MOLECULAR CHARACTERIZATION OF HEXOSE TRANSPORTER IN SUGAR ACCUMULATION OF PAPAYA FRUIT DURING MATURATION AND RIPENING

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI'I IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

HORTICULTURE

AUGUST 2007

By

Peerapong Sangwanangkul

Dissertation Committee:

Robert E. Paull, Chairperson
Kent Kobayashi
Paul Moore
Richard M. Manshardt
Monto Kumagai

ACKNOWLEDGEMENTS

I express my profound respect and gratitude to Dr. Jingtair Siripanich for giving me many opportunities and for recommending me for the doctoral position at the Department of Tropical Plant and Soil Sciences (TPSS). I convey my deepest gratitude to Prof. Dr. Robert E. Paull, my supervisor, for his guidance and support. He suggested the idea behind this work and gave me an opportunity to work on it. I express my sincere gratitude to Prof. Dr. Richard M. Manshardt, Prof. Dr. Kent Kobayashi, Dr. Paul Moore and Prof. Dr. Monto Kumagai, my dissertation committee, for their valuable suggestions and discussions.

I express my sincere appreciation to Dr. Helen Turano, Dr. Nancy Chen, and Ms. Beth Ishikura for their invaluable suggestions and technical assistance in all parts of the process, even those beyond work.

I express my special thanks to the staff at the Poamoho Experimental Research Station for management of the papaya field for this experiment and to all the TPSS secretaries for their kind and helpful support in all administrative activities.

My sincerest thanks to Mr. Roland J. Yap and his family for their kindness, aid, and friendship throughout my study, and for being my second family. They made me feel at home while I was far away from Thailand.

I deeply appreciate and thank Mr. James Keach, my good friend and roommate. He has always been warm, patient and very supportive. He also helped me improve my English.

I especially thank my colleagues at the Postharvest Research Unit in Kasetsart University and all friends, both in Thailand and the USA, for their friendship and support.

I sincerely thank the Postgraduate Education and Research Development Project in Postharvest Technology at Kasetsart University, Thailand, for financial support of my Ph.D.

I cannot verbalize the gratefulness I feel toward my parents, Ms. Mali & Mr. Wannasit Sangwanangkul, as well as my sisters and brothers for their care, affection, encouragement and support.

ABSTRACT

Papaya glucose uptake during fruit development was studied by comparing ^{14}C-glucose uptake of mesocarp slices between two cultivars, 'Sunset' and UH801 (low sugar line) in two seasons, September-December, 2003 and April-July, 2004. 'Sunset' papaya fruit, though one third the weight of UH801, had higher total soluble solids (TSS), total sugar, and dry weight (DW) than UH801. In 2003, both cultivars reached the color break stage at the same time, 125 days after anthesis (DAA), but flesh color, TSS, and DW of 'Sunset' fruit began to increase 111 DAA, one to three weeks before UH801. In 2004, both cultivars started to develop skin and fresh color and accumulate TSS, total sugar, and DW at 139 DAA. In 2003, glucose uptake in 'Sunset' fruit was lower than for UH801, but found to be higher in 2004. The pattern of ^{14}C-glucose uptake by 'Sunset' papaya flesh discs was divided into three fruit growth phases during fruit maturation. Uptake was initially low during the first, 90-97 DAA (2003) and 90-111DAA (2004), and last phase, after 132 DAA for both years and high 97-132 DAA and 11-132 DAA in 2003 and 2004, respectively. The maximum glucose uptake in 'Sunset' was found at 118 DAA in both years to be 1,367 and 2,140 μmol mg protein^{-1}h^{-1}, respectively. Papaya hexose transporter appeared to be an energy-dependent cotransporter. Hexose transporter activity was detected in both varieties but ^{14}C-glucose uptake did not appear to correlate with papaya sugar accumulation. Hence, the role of invertase in phloem unloading may be a more important factor in determining fruit sugar levels at harvest.

The 1,642 bp cDNA of the first *Carica papaya* hexose transporter (*CpHT1*) has been cloned. The sequence of *CpHT1* cDNA matched part of the sequence on the papaya genome supercontig_1226 which consisted of 3,218 bp with four exons and three introns. The full length *CpHT1* mRNA was predicted to be 1,732 bp and encoded a 523 amino acids long peptide. The predicted polypeptide was estimated to be 57.48 kDa and contained 12 transmembrane helices with both amino and carboxyl terminals located in the cytosol.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Acid invertase
bp	Base pair
BSA	Bovine serum albumin
cDNA	Complementary DNA
CIP	Calf intestinal phosphatase
DAA	Days after anthesis
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EB	Erythrosin B
EDTA	(Ethylenedinitro) - Tetracetic acid
GSP	Gene specific primer
Hyb	Hybridization
kDa	Kilo Dalton
LB	Luria-Bertani
MES	2-(N-morpholino) ethansulfonic acid
mRNA	Messenger Ribonucleic acid
PVP	Polyvinylpyrrolidone
RNA	Ribonucleic acid
RNase	Ribonuclease
SE/CCC	Sieve element / companion cell complex
SPS	Sucrose phosphate synthase
SS	Sucrose synthase
TAP	Tobacco acid pyrophosphatase
TMHs	Transmembrane helices
TSS	Total soluble solids

CHAPTER 1

INTRODUCTION

Sugar accumulation in fruit flesh is important for fruit growth and quality. In Hawaii, there is a minimal grade standard requirement of 11.5% soluble solids content (SSC) in papaya (Anon, 1990). To meet this standard papaya needs to be harvested when the fruit skin color develops at least 5% skin yellowing (Akamine and Goo, 1971). Immature fruit are lower in sugar content and ripen poorly. However, mature papaya fruit when ripe sometimes have low sugar content and poor flesh color development which lower market quality. Unlike many other climacteric fruits, papaya lacks stored carbohydrate reserves such as starch found in banana and mango, therefore papaya fruit must remain attached to the plant to accumulate more sugar. Papaya accumulates sugars mainly in the form of sucrose.

Sucrose is moved as a photoassimilate through the phloem from the leaf source to the sink organs. At the sink (fruit, root and shoot), sucrose is unloaded either via the apoplastic or symplastic pathways (Williams *et al.*, 2000). Symplastic unloading is thought to occur in growing tissue where the sucrose gradient is maintained by incorporation of the sucrose into cellular structures. In apoplastic unloading, sucrose is inverted by invertase, and the hexose sugars formed pass through the plasma membrane and tonoplast into the cytoplasm and the vacuole, respectively, via hexose transporters. The mechanism of hexose transport at the molecular level has been investigated in many plants (Ruan *et al.*, 1997; Shiratake *et al.*, 1997; Weber *et al.*, 1997; Fillion *et al.*, 1999; Gear *et al.*, 2000). The apoplastic pathway seems to occur late in fruit development in pear (Shiratake *et al.*, 1997), tomato (Ruan *et al.*, 1997; Gear *et al.*, 2000), grape berry (Fillion *et al.*, 1999), Fava bean seed (Weber *et al.*, 1997) and papaya (Zhou *et al.*, 2003).

Zhou and Paull (2001) found that sugar accumulation in "Sunset" papaya fruit flesh increased during the period 100 to 140 days after anthesis (DAA). This increase occurred after seed maturation and was related to an increase in acid invertase (AI) activity. AI activity began at 90 DAA and had the highest activity at 125 DAA whereas sucrose phosphate synthase and sucrose synthase activities, the other two main enzymes involved in sucrose metabolism,

remained low during this maturation stage though they were high during early fruit growth. The characterization of AI at the molecular level has also been reported (Zhou *et al.*, 2003).

Sugar accumulation in papaya fruit via the apoplast would require the presence of hexose transporter(s), however other than the stored sugar, there is no evidence that these transporters occur in papaya during sugar accumulation. Hexose transporter gene expression and hexose transporter activity have not been reported for papaya fruit. The absence of this data means that the importance of apoplastic unloading during sugar accumulation and the role of invertase is unsupported.

Objectives of this study are to determine whether hexose transporter genes are being expressed and the activity of hexose transporter during papaya fruit development, maturation, and ripening.

CHAPTER 2

LITERATURE REVIEW

2.1 Papaya fruit development, maturation and ripening

2.1.1 Introduction

2.1.1.1 Botany and fruit morphology

Papaya (*Carica papaya* L.) is a member of the family *Caricaceae* and is the only species in the genus of *Carica* (Badillo, 2000). It is a perennial herbaceous dicotyledonous plant which bears flowers and fruit for most of the year. The plants can be dioecious, producing male and female flowers on separate plants; monoecious, with male and female flowers on the same plant; or gynodioecious, with male and female reproductive organs in the same flower, called hermaphrodites. Plants are identical in appearance and the sex type can not be identified until the flowers appear, about 6 months after germination. Only female flowers and hermaphrodite flowers produce fruit (Muda *et al.*, 1994; Nakasone and Paull, 1998). Female plants are stable and produce female flowers only, whereas the flower types of the hermaphroditic plants vary depending on the seasonal temperature. During hot summer season (> 35°C), most hermaphroditic plants tend to form male flowers with non-functional female parts. In contrast, hermaphroditic papayas grown during low temperature (< 17°C) of winter season mostly produce female flowers with the male parts fusing with the carpel to form carpelloidy misshapen fruit (Awada, 1958; Hamilton, 1987; Nakasone and Paull, 1998)

Papaya plants generate the first flower about 6 months after germination (Muda *et al.*, 1994). In Hawaii, Solo papaya fruit development takes 150 to 164 days after anthesis to mature to full size and be ready to harvest. This period can be 2-3 weeks longer during the cold season (Paull, 1993). Papaya fruit shape can be spherical or oblong depending on the type of the flowers, female or hermaphrodite, respectively. Fruit size ranges from 5 to 40 cm long and can weight up to over 10 kg. Each fruit is composed of five carpels with a large central seed cavity (Paull, 1993). Papaya seeds are attached to the placenta in the large central cavity. The white

immature young seeds turn dark gray and then black during maturation. The immature white edible flesh (mesocarp) turns to orange-yellow, salmon-pink or red during ripening depending upon cultivar, whereas the fruit skin changes from green to greenish yellow, called color break stage, which is used as an harvesting index, then yellow when ripe (Paull, 1993; Nakasone and Paull, 1998).

2.1.1.2 Importance

Papaya is one of the major tropical fruit crops of Hawaii. The Solo papaya cultivar originated from Barbados and was initially introduced to Hawali in 1911 (Yee *et al.*, 1970). This cultivar became the parent line of many commercial cultivars developed in Hawaii, such as Kapoho, Sunrise, Sunset and Waimanalo (Nakasone and Paull, 1998). In Hawaii, the new commercial transgenic ringspot virus-resistant cultivars, Rainbow and Sunup, are of the Solo type (Gonsalves, 2004). In 2002, papaya fruit production in Hawaii totaled about 45.9 million pounds worth 11.9 million dollars. Ninety-three percent of the total production is utilized for the fresh fruit market and about 50% of this fresh fruit is exported (Hawaii Agricultural Statistics Service, 2004).

Papaya fruit is popular not only because of its good taste, but also because of its high nutritional value. It is a cheap source of vitamins and minerals and contains a high carotene content, a precursor of vitamin A, that varies between 1,160-2,431 μg /100 g edible portion, vitamin C (69-71 mg /100 g), calcium (11-31 mg /100 g), and potassium (39-337 mg /100 g) depending on the cultivar (Tee *et al.*, 1988).

Papain, a proteolytic enzyme that digests proteins is found in the milky latex of young green papaya fruit and is an important product from papaya. The enzyme is in a great demand in the United Kingdom and USA. Papain is used in meat tenderizing preparation, clarifying beer, in the manufacture of chewing gum and cosmetics, tanning hides, degumming of natural silk and to give shrink resistance to wool (textile and tanning industries) (Singh, 1980; Rohani, 1994; Nakasone and Paull, 1998).

2.1.2 Fruit development, maturation, and ripening

2.1.2.1 Fruit growth and development

Papaya fruit growth and development from pollination to maturation varies widely due to many factors including cultivars, age of bearing tree, season, and the selected stage as a maturity index (Nakasone, 1986). The growth pattern of 'Eksotika', 'Batu Arang' (Muda *et al.*, 1994), and 'Pakchong 1' (Kulwithit, 1993) papaya fruit is expressed as a simple sigmoid curve, which can be divided into three stages. The initial fruit growth stage is slow and associated with cell division. This stage is followed by a rapid increase in size and weight due to the cell enlargement. The third stage occurs as the fruit reaches full size and the seeds mature (Muda *et al.*, 1994). This development takes 140 days after anthesis in 'Eksotika', 'Batu Arang' (Muda *et al.*, 1994), and 'Pakchong 1' (Kulwithit, 1993) and 168 to 182 days in Hawaii's cultivars, with an extra two weeks delay during cool season (Nakasone, 1986).

2.1.2.2 Physico-chemical changes during fruit maturation and ripening

2.1.2.2.1 Color development

When papaya fruit approaches maturity, the seed color changes from white to gray and then black when fully mature. The edible flesh turns yellow, orange-yellow, salmon-pink or red during fruit maturation and ripening, depending upon cultivar. The fruit skin changes from green to greenish yellow, called color break stage, which is used as an harvesting index, then turns yellow at full ripe stage (Paull, 1993; Nakasone and Pauli, 1998). This skin color change initially occurs at the stigma end. The degree of skin color development has been widely used as a harvesting index (Pauli *et al.*, 1997).

2.1.2.2.2 Fruit softening

Papaya fruit softens as it matures and ripens. Papaya flesh begins to ripen from the inside out and takes around 6 to 12 days to reach an edible stage depending upon cultivar and when harvested at the color break stage (Pauli, 1993). In 'Pakchong 1' papaya, the fruit firmness slightly increased from 30 DAA until the fruit reached the maximum at 110 DAA of 95.1 Newton,

then decreased to 84.9, 82.2, 80.5 and 77.5 as fruit reaches the mature green stage, 5%, 15% and 25% skin yellow stages, respectively (Kulwithit, 1993). The decline in firmness has been associated with cell wall hydrolase activity.

The major cell wall degrading enzymes involved in papaya fruit softening are endoxylanase, polygalacturonase (PG), pectin methylesterase, β-1,4-D-glucanase and β-galactosidase. Xylanase and polygalacturonase activities are at their maximum when the fruit is 40-60% yellow, whereas the activity of pectin methyl esterase and glucanase are present at the beginning and continuously increase during ripening (Paull and Chen, 1983). Lazan *et al.* (1989) reported that β-galactosidase activity doubled during fruit ripening, but unlike PG activity, it was found consistently higher in the outer mesocarp than in the inner mesocarp. The role of β-galactosidase is unclear. It is thought that β-galactosidase contributes to pectin and hemicellulose modification by allowing greater access of hydrolases to the cell wall polysaccharides and thereby allows softening of the fruit (Lazan *et al.*, 1995)

2.1.2.2.3 Organic acids and voiatiles

The concentration of total organic acids and nonvolatile acids in papaya flesh decrease during fruit development and reach the minimum at the ripe stage (Selvaraj *et al.*, 1982). However, the titratable acidity of papaya flesh after harvest was found to increase only slightly during fruit ripening. The increase in acidity is believed to be associated with an increase in the level of free galacturonic acid (Lazan *et al.*, 1989; Paull, 1993). The three major organic acids in papaya fruit are citric, malic, and succinic acid (Ali *et al.*, 1994). During fruit ripening, citric acid content slightly decreases; malic acid content remains unchanged, whereas the concentration of succinic acid steadily increases (Ali *et al.*, 1994).

The sensory components of papaya fruit is based on the concentration of sugars and organic acids. The characteristic flavor of papaya fruit is mainly determined by the type and level of the volatile compounds present in the fruit tissue during ripening (Ali *et al.*, 1994). Up to 200 volatile compounds have been identified by the GC-MS in papaya (Flath and Forrey, 1977;

MacLeod and Pieris, 1983). However, the concentration of each volatile component is very low and varies among cultivars and location. Methyl benzoate is the only papaya volatile described as having papaya fruit quality on odor assessment (MacLeod and Pieris, 1983).

2.1.2.2.4 Sugar accumulation

Sugar accumulation is one of the more important biochemical changes of papaya fruit during development, maturation, and ripening. Since papaya does not store starch, the sweetness of papaya fruit after harvest relies on sugar accumulation during fruit development and maturation before harvest (Chen *et al.*, 2001). There are three major sugars present in papaya flesh: glucose, fructose, and sucrose (Sankat and Maharaj, 1997). In the early stage of papaya fruit development, glucose is the major sugar form and slowly increases until the fruit matures at 135 DAA, whereas the sucrose content is low in the young immature fruit and rapidly increase 110 DAA becoming the major sugar form, about 80% of the total sugar, in mature and ripe fruits (Chan *et al.*, 1979). Glucose, fructose, and sucrose content in the flesh do not significantly increase until 112 DAA. Sucrose rapidly increases one month before maturation or during the period of 100 to 140 DAA and accounts for 40-50% of total sugar in mature fruit (Zhou and Pauli, 2001). This increase occurs after seed maturation and is related to an increase in cell wall acid invertase (AI) activity (Zhou and Paull, 2001). The increase of AI activity parallels sugar accumulation during fruit maturation, suggesting that sugar unloading into the fruit mesocarp during this stage is apoplastic. The high levels of insoluble cell wall AI activity during this stage of increased sugar unloading, possibly maintains the sucrose concentration and pressure gradient between the source and the sink cells (Chen *et al.*, 2001) as found in tomato (Godt and Roitsch, 1997).

2.2 Factors affecting papaya fruit production and quality

2.2.1 Cultivars

Like other fruits, papaya fruit size and weight vary among cultivars. It can be as large as 40 cm and weight up to over 10 kg (Pauli, 1993). Cultivars and ripening stage of fruit at

harvest not only influence fruit size, they also affect final sugar composition of papaya (Akamine and Goo, 1971; Muda *et al.*, 1994) and skin and flesh color (Akamine and Goo, 1971; Ali *et al.*, 1994).

2.2.2 Temperature

Temperature of the growing season is one of the main factors affecting papaya fruit growth, development and quality. In subtropical area, papaya fruit set does not occur in winter. Moreover, fruit set before winter often delays maturity for about 90 days from the normal 120-150 days (Nakasone and Pauli, 1998). In Hawaii, the growth period of papaya is prolonged about 2 weeks during cool season. Fruit set of 'Sunset' papaya in June reaches the color break stage at 140 days after anthesis (DAA) while fruit of the same cultivar set in October of the same year required 180 DAA to reached color break stage (Nakasone, 1986).

Fruit that develop during winter also have lower total soluble solids. Low temperatures below 15°C during the early phase of fruit growth significantly delay the growth and reduce the fruit size, possibly due to the reduction of cell division (Nakasone and Pauli, 1998). In contrast, papaya fruit growing during warmer season has a longer fruit length (Ong, 1983).

2.2.3 Cultural practices

Cultural practices including plant spacing, irrigation, mulching, pruning, thinning, and fertilizers influence fruit quality (Pantastico, 1975). Fruit thinning results in a high percentage of well formed, uniform marketable sized fruit (Muda *et al.*, 1994). The study of source-sink relationships in 'Line 8' and 'Sunset' papaya showed that fruit thinning increases new fruit set and ripe fruit TSS level. Fruit thinning in 'Line 8' also increased TSS and sugar contents in young fruit and the remaining fruit on plants, compared to the same aged fruit on the control and defoliated plants (Zhou et al., 2000). Single defoliation significantly reduced new flower, fruit set and total soluble solids (TSS) in ripe papaya fruit, but had no effect on fruit production. Moreover, continual defoliation reduced fruit size, sugar content, cell wall acid invertase activity, and fruit production (Zhou et al., 2000). The number of fruit per leaf or crop load affects apple fruit quality and aroma

(Mattheis and Fellman, 1999). The influence of crop load on the impact on compounds contributing to fruit flavor and aroma of 'Jonagored' apples shows that apple grown with the lowest fruit:leaf ratio has the highest total soluble solids, titratable acidity, butyl and hexyl acetate contents, aroma compounds in apple fruit, throughout a late season harvest period (Poll *et al.*, 1996).

Irrigation increases both the number and size of marketable papaya fruit (Muda *et al.*, 1994), however, papaya growing in wet condition or having excessive irrigation tends to produce misshaped carpelloidic fruit (Muda *et al.*, 1994; Nakasone and Pauli, 1998), low sugar content and larger fruit size and increases disease problems (Awada and Ikeda, 1957; Muda *et al.*, 1994; Nakasone and Paull, 1998). Papaya growing under low soil moisture condition tends to shift the flower sex type to male, causing flowers and fruitlets abortion leading to fruiting skips along the trunk and lower fruit yield (Muda *et al.*, 1994; Nakasone and Paull, 1998) but the remaining fruit have increased fruit sugar content (Awada and Ikeda, 1957).

2.2.4 Fertilizers

During fruiting stage, potassium fertilization improves fruit quality by increasing fruit size and enhancing flesh color, sugar content, and fruit firmness (Nakasone, 1986; Muda *et al.*, 1994). Phosphorus levels should be limited during fruit development as high level is reported to reduce fruit size. Like banana, nitrogen fertilizer does not show any influences on the papaya fruit size and quality. The level of nitrogen requirement is similar from the juvenile to the fruiting stage (Muda *et al.*, 1994). Boron deficiency leads to low fruit set, produces malformed fruit with rough or 'bumpy' surfaces and the fruit often ripen unevenly and have low sugar content. Most seeds from boron deficient fruit are either aborted, poorly developed or absent (Wang and Ko, 1975; Chapman *et al.*, 1978).

2.2.5 Plant growth regulators

Shanmugavelu *et al.* (1973) reported that Ethephon, Alar and Phosfon D at 250 ppm applied to papaya increases the total and reducing sugar content of the fruit. Similarly, gibberellic

acid (GA) at 200 ppm increases TSS, fruit acidity and ascorbic acid content of 'Coorg Honey Dew' papaya (Shanmugavelu *et al.*, 1973). This response is different from that reported for GA-treated cherries where TSS was reduced (Modlibowska and Wickenden, 1982). Cheikh *et al.* (1992) also found that GA influenced the steady-state level of SPS and its activity in soybean and spinach.

2.3 Enzymes involved in sugar accumulation

Three major enzymes are involved in sugar metabolism: sucrose phosphate synthase (SPS), sucrose synthase (SS), and acid invertase (AI). Sucrose phosphate synthase (E.C. 2.4.1.14) is involved in sucrose synthesis from UDPG and fructose-6-phosphate (Bruneau *et al.*, 1991). Sucrose synthase (E.C.2.4.1.13) can function both in sucrose synthesis and cleavage (Sung *et al.*, 1988; Wang *et al.*, 1994). Acid and neutral invertases (β-fructofuranosidase, E.C.3.2.1.26), catalyze the hydrolysis of sucrose into glucose and fructose. Cell wall bound acid invertase is mostly active in apoplastic sinks such as developing seed (Weber *et al.*, 1995, Weschke *et al.*, 2003), and fruit (Godt and Roitsch, 1997; Ruan and Patrick, 1995; Zhang *et al.*, 2001; Miron *et al.*, 2002). Cell wall AI hydrolyzes phloem-unloaded sucrose into glucose and fructose which in turn are transported across plasma membrane via hexose transporter, a sink-specific sugar transporter (Figure 2.1) (Sauer *et al.*, 1994; Williams *et al.*, 2000). The cleavage of sucrose by AI allows the sucrose concentrationin the phloem from the leaves to the fruit to be maintained. Due to the location of the AI in the cell wall and the hexose transporter on the plasma membrane, and the relationship between AI activity and sugar accumulation, plant physiologists recognize that there is cooperation between AI and the hexose transporters as a fact. There is some evidence indicating an increase in AI activity correlates with the increase in sucrose accumulation during fruit maturation and ripening of papaya (Chan and Kwok, 1976; Zhou and Pauli, 2001), tomato (Ruan and Patrick, 1995; Ruan *et al.*, 1997), muskmelon (McCollum *et al.*, 1988), and apple (Zhang *et al.*, 2001). The relationship between AI activity and sugar accumulation supports the concept of apoplastic unloading. Further evidence that supports apoplastic unloading is the location and amount of AI and plasmodesmata.

Immunolocalization of acid invertase in apple fruit mesocarp shows that AI is mainly located on the cell walls of the sieve element/companion cell complex (SE/CCC), phloem parenchyma cells and other parenchyma cells. The number of AI (the density of immunogold particles) increased during fruit development and this related to the increasing in its activity (Zhang *et al.*, 2001). Numerous plasmodesmata have been found between apple parenchyma cells, but almost no plasmodesma have been found between SE/CCC and the surrounding parenchyma cells. This evidence confirms apoplastic phloem unloading in developing fruit tissue (Zhang *et al.*, 2001).

In the apoplast, sucrose is hydrolyzed by AI to hexoses, which are transported into the cytosol via energy-dependent plasma membrane hexose transporters, followed by the resynthesis of hexoses into sucrose by SPS in the cytosol (Ruan and Patrick, 1995; Brown *et al.*, 1997; Ruan *et al.*, 1997; Miron *et al.*, 2002). The increase in AI activity in sucrose hydrolysis at the unloading end of the phloem (sink site), is believed to maintain or increase the sucrose gradient, leading to a faster unloading of sucrose from the SE/CCC into the sink cells (Weber *et al.*, 1995; Patrick, 1997).

In 'Sunset' papaya fruit (Zhou and Pauli, 2001), sugar accumulation increased during the period 100 to 140 days after anthesis (DAA). This increase occurred after seed maturation and was related to an increase in AI activity. AI activity began at 90 DAA and had the highest activity at 125 DAA whereas SPS and SS activities, the other two main enzymes involved in sucrose metabolism, remained low during this maturation stage, though they were high during early fruit growth. SS is a major sugar hydrolysis enzyme that was found correlated with the fruit

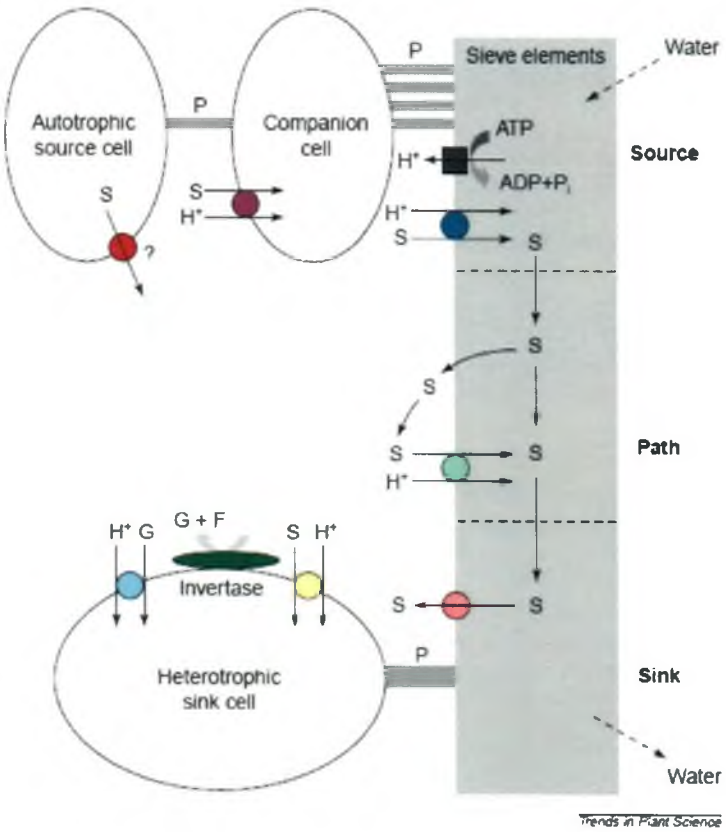

Figure 2.1. Sugar loading and unloading between source and sink. Sucrose (S) in the mesophyll cell of the source leaves symplastically loads into the sieve elements-companion cell complex from cell to cell via plasmodesmata (P). In many species, sucrose escapes the symplast at some points, possibly via sucrose carrier (red dot) and then actively transports from the apoplast (intercellular space) into the companion cell and/or sieve elements via plasma membrane H^+ /sucrose symporters (purple and dark-blue dots, respectively). The proton was generated from the plasma membrane H^+ - ATPase (black square). Once sucrose is loaded into the sieve elements of the phloem, it is transported in the phloem sap by bulk flow. However, if sucrose leaks out from the phloem, it can be moved back via the H^+ /sucrose symporters (green dot) along the path. At the sink cell, sucrose can be symplastically unloaded from the sieve elements into the sink cell via plasmodesmata or be transported into the apoplast via sucrose carrier (pink dot). In the apoplast, sucrose can be either directly moved across plasma membrane sucrose transporter (yellow dot) or first inverted into glucose (G) and fructose (F) by the cell wall acid invertase (dark-green ellipse) and then moved across plasma membrane hexose transporter (blue dot) into the cytosol (From Williams *et al.*, 2000).

respiration and carbon import rate during early fruit growth (Zhou and Paull, 2001). This result

respiration and carbon import rate during early fruit growth (Zhou and Paull, 2001). This result confirms the earlier reports of the high activity of the soluble AI (Chan and Kwok, 1976) and the low level of the SPS and SS activities during papaya fruit ripening (Hubbard *et al.*, 1991).

Like pineapple fruit (non-climacteric fruit), SS plays a major role in sink establishment during the early papaya fruit development. The increase in AI activity plays a major role in apoplastic phloem unloading during fruit maturation and sugar accumulation just before ripening and might be responsible for maintaining pressure gradient between source and sink cell resulting in the final fruit sweetness (Chen *et al.*, 2001) as reported in tomato during late fruit development (Ruan and Patrick, 1995; Godt and Roitsch, 1997).

The characterization of papaya cell wall AI at the molecular level has also been reported (Zhou *et al.*, 2003). The complete amino acid sequence of papaya AI has an open reading frame that encodes a predicted polypeptide chain of 582 residues and calculated molecular weight of 65,537 Da and is 68% and 45% identical with carrot apoplastic and vacuolar invertases. Its activity is regulated at both the transcriptional and translational levels (Zhou *et al.*, 2003).

2.4 Source-sink relationship and sucrose transport during fruit growth, development and maturation

2.4.1 Source-sink reiationship

Source-sink relationship plays a major role in fruit and seed development and especially in sugar composition and accumulation during fruit maturation and ripening as found in peach (Pavel and DeJong, 1993) and papaya (Zhou *et al.*, 2000).

Sugar accumulation in fruit flesh is important for fruit growth and quality (Pavel and DeJong, 1993; Zhou *et al.*, 2000). The optimum number of leaves and their area for development of individual fruit has been reported for kiwifruit (Antognozzi *et al.*, 1992), mango (Chacko *et al.*, 1982), apple (Palmer *et al.*, 1991) and papaya (Zhou *et al.*, 2000). Unlike most fruit trees, papaya is an indeterminate plant that develops new leaves and fruit and has fruit at all stages of development present on a single plant (Nakasone, 1986). Once the initial flowering occurs,

papaya tree flowers and sets continuously, requiring an uninterrupted carbohydrate supply for fruit development and quality. The competition between vegetative and reproductive sinks and between young and mature fruit sinks exist through all stages of plant development after the juvenile phase (DeJong *et al.*, 1987; Jeuffroy and Warembourg, 1991). Study of the relationship between leaf area, fruit production and sweetness in papaya by Zhou *et al.* (2000) showed that 75% defoliation significantly reduced new flower production, fruit set, and TSS in the ripe fruit, whereas 50% defoliation had no effect. Fruit thinning increased new fruit set, size, growth rate, sugar content and AI activity in the immature young fruit and TSS in the ripe fruit after old fruit removal. Each papaya mature leaf was able to support about three developing fruit.

2.4.2 Phloem structure, loading pathway and long distance transport of sugar

Most higher plants convert fixed CO_2 into sucrose, a non-reducing sugar, which serves as the major sugar transported in the phloem (Ziegler, 1975; King *et al.*, 1997). Photoassimilate (sucrose) is moved through the phloem sieve elements from the source to the sink organs. Since most organelles of the phloem sieve elements (including vacuoles, ribosomes and nuclei) are degraded during sieve element development, differentiation, and longitudinal extension (Esau and Gill, 1971, 1972; Thorsch and Esau, 1981; Eleftheriou, 1987; Sjölund, 1997; Taiz and Zeiger, 2002), the sieve elements exist as a complex with the adjacent companion cells, the so called sieve element-companion cells complex (SE-CCC). The companion cells provide energy and proteins to the sieve elements via numerous plasmodesmata (Fisher et al., 1992; Ishiwatari et al., 1995; Taiz and Zeiger, 2002). Once sucrose from the source cell moves into SE-CCC, the accumulation of sucrose in the SE-CCC induces the osmotic uptake of water forcing the phloem sap to flow along the sieve tube (Figure 2.1) from the source end to the sink end by bulk flow (Williams *et al.*, 2000). The pressure gradient is maintained by the loss of water from the sieve tube during unloading of sugar at the sink cell (Williams *et al.*, 2000; Taiz and Zeiger, 2002).

2.4.3 Phloem unloading pathway

At the sink (fruit, root, and shoot), sucrose is unloaded either via the symplastic or

apoplastic pathway (Figure 2.1) depending on the species, organ or tissue, and development stage (Oparka, 1990; Taiz and Zeiger, 2002). Both pathways might operate in the same organism, organ or tissue. Some sinks such as young leaves and root tips can import photoassimilates directly from the phloem via numerous plasmodesmata whereas other sinks such as pollen grains and tubes, developing embryo, guard cells are not connected via plasmodesmata to the phloem (Thorne, 1985; Patrick, 1990; Patrick *et al.*, 1995; Taiz and Zeiger, 2002; Stadler *et al.*, 2005a). For these tissues without plasmodesmata, importation of photoassimilates into these cells and vacuoles occurs only via transport proteins in their plasma membranes and tonoplast by the so-called apoplastic pathway (Imlau *et al.*, 1999; Büttner and Sauer, 2000; Williams *et al.*, 2000). Unlike source-leaves, sinks vary widely in structure and function from growing vegetative organs such as root tips and young leaves to storage tissues like roots and stems to reproductive organs like fruits and seeds, there is no single scheme of phloem unloading. Some sinks depend entirely on symplastic phloem unloading, while others prefer entirely or partially apoplastic phloem unloading (Oparka and van Bel, 1992; Taiz and Zeiger, 2002).

2.4.3.1 Symplastic unloading pathway

Symplastic unloading is thought to occur in growing tissue where the sucrose gradient is maintained by incorporation of the sucrose into cellular structures. The utilization of transport sugar in the growing tissues such as a substrate for respiration, and for the synthesis of the other molecules needed for growth and storage polymer formation, maintains a low sucrose concentration in the sink cells, thus maintaining a concentration gradient for sugar uptake (Patrick, 1990). Transport of sugar through plasmodesmata occurs passively from a high sucrose concentration in the sieve elements to a low sucrose concentration in the sink cells. Meristematic and elongating regions of the primary root tips contain large numbers of plasmodesmata sufficient for symplastic phloem unloading (Schulz, 1995; Stadler *et al.*, 2005b). In some young dicot leaves such as sugar beets (Geiger *et al.*, 1973; Gougler Schmalstig and Geiger, 1985) and tobacco (Ding *et al.*, 1988; Oparka *et al.*, 1999), the phloem unloading appears to be completely

symplastic movement. The study of the insensitivity to the *p*-chloromercuribenzenesulfonic acid (PCMBS), a reagent that inhibits active transport of sucrose across plasma membranes, is provided as evidence for the symplastic phloem unloading in these sink tissues (Gougler Schmalstig and Geiger, 1985; Patrick, 1990; Taiz and Zeiger, 2002).

2.4.3.2 Apoplastic unloading pathway

In apoplastic transport, sink ceils can either actively import sucrose from the apoplast directly via sucrose transporters or, alternatively, sucrose can be inverted by cell wall bound invertase into glucose and fructose which are in turn taken up via plasma membrane and tonoplast hexose transporters into the cytoplasm and the vacuole, respectively (Imlau *et al.*, 1999; Büttner and Sauer, 2000; Williams *et al*, 2000).

The apoplastic pathway seems to occur late in fruit development in pear (Shiratake *et al.*, 1997), tomato (Ruan and Patrick, 1995; Godt and Roitsch, 1997; Ruan *et al.*, 1997; Gear *et al.*, 2000), grape berry (Fillion *et al.*, 1999), and Fava bean seed (Weber *et al.*, 1997) and possible papaya (Zhou and Pauli, 2001). Although, the sucrose uptake in young fruit is characterized by symplastic pathway as found in tomato (Ruan and Patrick, 1995) and grape (Zhang *et al.*, 2006), the uptake of sucrose is apoplastic when the fruit start to mature and ripen (Ruan and Patrick, 1995; Zhang *et al.*, 2006). In grape berry, the turning point of phloem unloading pathway from symplastic to apoplastic pathway occurs at or just before the onset of fruit ripening (Zhang et al., 2006). The transition from symplastic unloading in young fruit to apoplastic unloading in developing and mature fruit possibly occurs in papaya (Zhou and Pauli, 2001). Zhou and Paull (2001) reported a correlation between an increase in AI activity and sugar accumulation during papaya fruit maturation and proposed that sugar unloading in papaya fruit during maturation is possibly via the apoplastic pathway, similar to the hypothesis of Hubbard *et al.* (1991). Hubbard *et al.* (1991) hypothesized that sucrose and invertase may be present in either or both different intracellular locations or different cells within the papaya mesocarp tissues. They suggested that the sucrose is unloaded from the phloem into the apoplastic pathway via a channel or transporter. A high level of SS activity occurs in young immature papaya fruit that

declines throughout growth and maturation, whereas AI activity is low in young fruit and increases when the fruit starts to mature. The increase in AI correlates with the increase in sugar accumulation (Zhou and Pauli, 2001). The relationship in papaya between enzyme activities and sugar accumulation during fruit maturation and ripening indicates a transition of the sugar transport pathway from symplastic to apoplastic (Zhou and Pauli, 2001). Therefore, if sugar unloading is apoplastic and acid invertase involves in sugar accumulation, then a hexose transporter must be involved (Williams *et al.*, 2000). The mechanism of hexose transport at the molecular level has been investigated in many plants (Ruan *et al.*, 1997; Shiratake *et al.*, 1997; Weber *et al.*, 1997; Fillion *et al.*, 1999; Gear *et al.*, 2000) and should be possible also for papaya.

The mechanism of sugar accumulation in papaya fruit via the apoplast requires the action of hexose transporter(s), however, there is no evidence that these transporters occur in papaya during sugar accumulation. The existences of hexose transporter gene expression and hexose transporter activity have not been reported for papaya fruit. Absence of such data means that the importance of apoplastic unloading during sugar accumulation and the role of invertase is unsupported.

2.4.4 Structure and functions of hexose transporter protein and genes among plant species

2.4.4.1 Hexose transporter superfamily

All plant hexose transporters belong to a large multigene superfamily of transmembrane facilitators (MFS, major facilitator superfamily) (Marger and Saier, 1993) and consist of at least seven members in *Chenopodium rubrum* (Roitsch and Tanner, 1994), eight members in castor bean (*Ricinus communis*) (Weig *et al.*, 1994), three genes in tomato berry (Gear *et al.*, 2000) and at least 14 genes encoding putative monosaccharide transport proteins (MST) in *Arabidopsis thaliana* (Büttner *et al.*, 2000; Williams *et al.*, 2000). The transporters have been described from the plasma membrane and tonoplast (Rausch, 1991; Shiratake *et al.*, 1997). Hexose transporters are generally sink-specific and are necessary for the import of glucose and fructose into sink cells after phloem-unloaded, sucrose is hydrolyzed by cell wall bound

invertases (Sauer *et al.*, 1994).

Hexose transporters have been isolated from several plants including *Arabidopsis*, tobacco, grape berry, and tomato (Sauer *et al.*, 1990; Sauer and Stadler, 1993; Fillion *et al.*, 1999; Gear *et al.*, 2000). Unlike disaccharide transporter which appears to be specific for plants, the monosaccharide transporters in higher plants are homologous to the monosaccharide transporters found in yeast, bacteria, fungi, blue-green algae, and mammalians including humans (Sauer and Stadler, 1993; Williams *et al.*, 2000).

Among the monosaccharide transporter genes in *Arabidopsis*, *AtSTP1* mRNA is most abundant and strongly expressed in source leaves and is also found in other organs including stem, flowers and roots (Sauer *et al.*, 1990). Most of the other *AtSTPs* and other transporter clones are highly expressed in sink organs such as *AtSTP2* in developing pollen (Truernit *et al.*, 1999), *AtSTP3* in leaves and sepals (Tanner and Caspari, 1996), and *AtSTP4* in wounded leaves (Truernit et al., 1996), anthers and root tips (Truernit *et al.*, 1999). In tobacco, *NtMST1* is expressed in roots, flowers, and young leaves (Sauer and Stadler, 1993). In tomato, both *LeHT1* and *LeHT3* are highly expressed in young fruit and root tips while *LeHT2* shows a high level of expression in source leaves and flowers (Gear *et al.*, 2000). The phylogenetic analysis indicates these tomato genes fall into 3 distinct subclasses out of the at least 4 subclasses of hexose transporters identified in higher plants.

2.4.4.2 Structure and functions of hexose transporters in diverse locations

Both monosaccharide and disaccharide transporters have typically 12 transmembrane-spanning domains structure. This transmembrane domain is a typical feature of all members of the MFS family (Marger and Saier, 1993). The homology between the amino-terminal halves and the carboxyl-terminal halves of these transport proteins indicating that they may have evolved by gene duplication from an ancestral gene coding for a six-transmembrane helix transporter (Griffith *et al.*, 1992; Marger and Saier, 1993; Sauer and Tanner, 1993). Until now, all of the monosaccharide transporters are reported to be energy-dependent H^+-symporters and accept both hexoses and pentoses that form a pyranose ring (Büttner and Sauer, 2000).

Functional analysis of plant monosaccharide transporters in heterologous expression systems, such as yeast and *Xenopus* oocytes, shows that the substrate specificities of the plant monosaccharide transporters is relatively broader than transporters from bacteria.

Each transporter preferably transports a different monosaccharide and/or occurs in different tissues. The relative transport rates of each transporter for different sugar substrates such as D-fructose, D-glucose, D-galactose, D-xylose, and D-mannose are dramatically different. In lower plants like *Chlorella sp.*, the monosaccharide transporters *CkHUP1* and *CkHUP3* preferably transported D-glucose, whereas *CkHUP2* transports D-galactose better than D-glucose (Sauer *et al.*, 1990a; Stadler *et al.*, 1995). In *Arabidopsis,* the sugar transporter AtSTPs has low specificity and has distinct functions in many different plant tissues (Sauer *et al.*, 1990b; Boorer *et al.*, 1994; Stolz *et al.*, 1994; Truernit *et al.*, 1996 and 1999; Sherson et al., 2000). The functional characteristics of *AtSTP1* gene has been studied in *Xenopus oocytes* and *Saccharomyces cerevisiae,* and shown to be a high affinity monosaccharide-H^+symporter capable of transporting several monosaccharides except fructose (Boorer *et al.*, 1994; Stolz et al., 1994).

Sherson et al. (2000) reported the expression of *AtSTP1* gene in germinating seeds and seedlings, mainly in the seedling root. *AtSTP1* is active before seed germination and is the major monosaccharide transporter in *Arabidopsis* seedlings that mainly transport D-galactose and D-mannose, but not D-fructose and L-arabinose. Comparing the growth of an *Atstp1* knock-out mutant and a wild type on high D-galactose content media, which inhibits growth of the wild type plant, has shown that the *Atstp1* mutants grow effectively even at up to 100 mM D-galactose. This result clearly indicates that the active transport by *AtSTP1* plays a major role in transporting exogenous sugar at a very high concentration to the embryo and the seedlings (Sherson *et al.*, 2000). *AtSTP1* peptide sequence is predicted to be 522 amino acids with a molecular weight of 57.518 kDa (Sauer *et al*, 1990b). The *AtSTP2* gene is found to be expressed in developing pollen where it possibly functions in the uptake of glucose derived from the degradation of the callose walls that surround the pollen tetrads during pollen maturation (Truernit *et al.*, 1999). The suggestion was that callose wall carbohydrates serve as a carbon source for early pollen

development (Truernit *et al.*, 1999). Expression analysis of *AtSTP4* shows that it occurs in the specific tissues at the very tip of root and pollen grains (Truernit *et al.*, 1996) similar to the expression of the tobacco monosaccharide transporter *NtMST1* in roots (Sauer and Stadler, 1993). *AtSTP4* is also expressed in pathogen infected tissues and wounded leaves. The increase in amount of *AtSTP4* in wounded leaves suggests a role in response to pathogen attack (Truernit *et al.*, 1996), similar to *AtSTP3* that is expressed in wounded leaves (Büttner *et al.*, 2000).

The role of monosaccharide transporters in developing seed has been studied in some plants. In Faba bean seed (*Vicia faba*), *VfSTP1* was found to be expressed in the embryo cotyledon epidermal cells (Weber *et al.*, 1997). Apparently, some of these monosaccharide transporters are expressed in various organs of plant such as *AtSTP1* expressed in leaves, stems, flowers, and roots (Sauer *et al.*, 1990b), whereas others are cell specific like *AtSTP2* in developing pollen (Truernit *et al.*, 1999). In addition, these transporters may play a role in defense reactions such as *AtSTP4* which is strongly induced by wounding and pathogen attack (Truernit *et al.*, 1996). Most of the genes are found expressed in the sink tissues that are depend on the imported photoassimilates from the source tissue like green leaves, or expressed following increased cellular metabolism after wounding. Büttner and Sauer (2000) concluded that monosaccharide transporters play an important role for the supply of carbohydrates to the non-green, rapidly growing or metabolically hyperactive tissues.

Results obtained from the kinetic study of the monosaccharide-H^+symporter, *AtSTP1* suggested that sugar uptake by *AtSTP1* is a 'sequential' mechanism in which protons and sugar molecules are imported in two ordered, sequential steps (Boorer et al., 1994). This sugar uptake mechanism in plants differs from the mechanism of the Na^+-dependent mammalian glucose symporter that import its substrates in a 'simultaneous' mechanism, substrate (sugar) and cosubstrate (Na^+) are transported together (Parent et al., 1992a, 1992b). In *Chlorella*, the monosaccharide transporter *CkHUP1* also preferred a 'simultaneous' transport mechanism (Komor and Tanner, 1974).

2.5 Summary

Sugar accumulation is influenced by genotype, temperature, cultural practices, fertilizers, and plant growth regulators. Three major enzymes are involved in sugar metabolism: SS, SPS, and AI. In papaya, AI activity increases during late fruit development, maturation and ripening and correlates with the increase in sugar accumulation. This increase in AI activity suggests that sugar unloading during fruit maturation is possibly through an apoplastic pathway. Apoplastic sugar unloading requires the action of a hexose transporter on the plasma membrane. Up until now, at least fourteen hexose transporters have been reported expressed in various tissues of *Arabidopsis* plants. At least three hexose transporters are found in tomato fruit, seven in grape berry and two in peach fruit. However, there is no report of papaya hexose transporters. The function, properties, structure and expression of this hexose transporter gene and protein are unknown. Hence, a study on the relationship between sugar accumulation and the activity of hexose transporter and its gene expression and structure during fruit development, maturation, and ripening could lead to an improvement in papaya fruit sugar content and an understanding of how fruit sugar content is regulated.

CHAPTER 3

HYPOTHESES AND APPROACHES

3.1 Hypotheses

The mechanism of sugar accumulation in papaya fruit via the apoplast requires the action of hexose transporter(s), however, there is no evidence that these transporters occur in papaya fruit during sugar accumulation. Absence of this data means that the importance of apoplastic unloading during sugar accumulation and the role of invertase is unsupported.

The hypotheses were:

I. Hexose transport across the plasma membrane and the tonoplast, and hexose transporter gene expression correlate to sugar accumulation during fruit development, maturation and ripening.

II. The hexose transporter activity and gene expression occur simultaneous with cell wall acid invertase activity and gene expression.

3.2 Approaches

I. To compare the sugar accumulation behaviors between two papaya cultivars differing by having high or low sugar accumulation.

II. To determine the relationship between sugar accumulation and hexose transporter gene expression and activity during fruit development, maturation, and ripening.

III. To determine the hexose transporter gene structure and its expression.

CHAPTER 4

PAPAYA FRUIT DEVELOPMENT AND C^{14}-GLUCOSE UPTAKE DURING FRUIT MATURATION AND RIPENING

4.1 Introduction

Papaya fruit accumulates the non-reducing sugar, sucrose (Chan *et al.*, 1979). Sucrose is moved from the source through the phloem to the sink organs (fruit, root, and shoot) and then unloaded either by the apoplast or symplast pathways (Williams *et al.*, 2000). Symplastic unloading is thought to occur in growing tissues where the sucrose gradient is maintained by incorporation of the sucrose into cellular structures and for cell maintenance. In apoplastic unloading, sucrose is inverted into two hexose sugars by invertase. These hexose sugars are moved via hexose transporters on the plasma membrane and tonoplast to the cytoplasm and the vacuole, respectively. The apoplastic pathway occurs late in fruit development in pear (Shiratake *et al.*, 1997), tomato (Ruan *et al.*, 1997; Gear *et al.*, 2000), grape berry (Fillion *et al.*, 1999), Fava bean seed (Weber *et al.*, 1997) and papaya (Zhou and Paull, 2001).

Zhou and Paull (2001) reported that the increasing of sugar accumulation in 'Sunset' papaya fruit occurred after seed maturation and during iate fruit development (100-140 DAA). The accumulation of sucrose was related to an increase in acid invertase (AI) activity. AI activity begins 90 DAA, before sugar accumulation has started, and has the highest activity at 125 DAA. Activities of sucrose phosphate synthase and sucrose synthase, the other two main enzymes involved sucrose metabolism, remained low during this stage (Zhou and Paull, 2001). The increase in AI activity during fruit maturation suggests that sugar unloading in papaya fruit tissue is apoplastic. Sugar unloading via the apoplastic pathway requires energy-coupled hexose transporters on the plasma membrane and the tonoplast (Gear *et al.*, 2000). Plant hexose transporters require a proton (H^+) gradient generated by ATPase in order to function as H^+/hexose symporters (Bush, 1993). The effects of plasma membrane ATPase inhibitors such as erythrosine B (EB), 2,4-dinitrophenol (DNP), antimycin A, carbonyl cyanide *m*-chlorophenyl-

hydrazone (CCCP), *N*-ethylmaleimide (NEM), KNO_3, and *p*-chloromercuribenzene sulfonic acid (PCMBS) on the hexose transporter activity have been studied in many fruits including tomato (Gear *et al*, 2000), pear (Shitake *et al*, 1997) and even on the citrate uptake by tonoplast vesicles of acidless *Citrus* juice cells (Canel *et al.*, 1995). Shiratake *et al.* (1997) reported that nitrate (KNO_3) inhibited approximately 80% of proton transport activity and closely corresponded to ATPase activity. Sugar uptake by pear fruit tissues tonoplast vesicles using a membrane and gel filtration assay methods showed the highest uptake for glucose, then fructose, sucrose, and sorbitol. The uptake of these four sugars was significantly inhibited by PCMBS. Shriatake *et al.* (1997) concluded that the uptake of these sugars was mediated by a transporter.

The objectives of current study were to determine the relationship between fruit weight, skin, flesh and seed color development, sugar and dry mass accumulation and the activity of hexose transporter during fruit maturation and ripening. The pattern of glucose uptake during fruit maturation and ripening was compared between the low-sugar line, 'UH801', and the high-sugar cultivar Sunset.

4.2 Materials and methods

4.2.1 Plant Material

Papaya (*Carica papaya* L.) was grown at the University of Hawaii Poamoho Experimental Station on Oahu, Hawaii. Papaya flowers of selected plants were tagged every week from anthesis. Three papaya fruits of the cultivars Sunset and UH801 (a low sugar line), were harvested weekly after the seeds began to mature. Seed maturation occurred about 90 days after anthesis (DAA) (±3 days) and the full ripe stage about 160 DAA during September-December, 2003 and April-July, 2004. Mesocarp samples were taken from opposite sides of each individual fruit and immediately frozen in liquid nitrogen and stored at -80°C until used.

4.2.2 Fruit growth

Papaya fruit weights were recorded weekly and reported in grams per fruit (g $fruit^{-1}$).

4.2.3 Skin, flesh and seed color determination

Skin color development was subjectively determined by visual estimation of the area of skin yellowing (0-100% yellow). Flesh color development was subjectively estimated as the degree of internal flesh showing carotenoid development (0-100% orange/red). Seed color development was subjectively estimated from the degree of darkening and sacrotesta formation (0-100% black).

4.2.4 Firmness determination

Fruit firmness was determined as the force (Newtons) required to depress a 1.5 cm-disc 2 mm into the fruit using AccuForce CADET model LKG-14 (Ametek inc., Largo, Florida).

4.2.5 Total soluble solids (TSS)

The percentage of the soluble solids content (SSC) was measured by a digital handheld refractometer (Palm Abbe model AR200, MISCO, Cleveland, Ohio).

4.2.6 Total sugar content

Total sugar content (mg g^{-1} of fresh weight) was determined by the method of Dubois (1954). Two grams of fresh tissue from both sides of a fruit was ground in 2 ml of 90% ethanol. The homogenate was centrifuged at 8000 rpm for 15 min. The supernatant was diluted with deionized water and then mixed with 25µl of 80% (w/w) phenol. 2.5 ml of H_2SO_4 was added and mixed well. After standing at 25°C for 10 min, the mixture was mixed again and allowed to stand for further 30 min. The absorbance was measured at 485 nm. The concentration of total sugar in the sample was determined with D-glucose as standard. The result was expressed as total sugar content mg D-glucose g^{-1} of fresh weight.

4.2.7 Dry mass assay

The dry mass percentage of each sample was determined by drying ten grams of fruit flesh tissue at 60-70^{o}C for at least 72 hrs, or until the weight was stable.

4.2.8 Total protein determination

Total protein was extracted by the method of Egusa and Pauli (1991) and then determined by the method of Bradford (1976), using bovine serum albumin (BSA) as a standard. The total protein content was expressed as $\mu g\ g^{-1}$ fresh weight.

4.2.9 Study of ^{14}C-glucose transport uptake

The study of ^{14}C-glucose transport was modified from the method of Ruan and Patrick (1995). Discs of papaya mesocarp were sampled using a cork borer, 8.0 mm in diameter from opposite side of each fruit (three fruit per cultivar) harvested each week during the period 90-160 days after anthesis (DAA). Each mesocarp disc was cut transversely into ten slices, 0.1 cm thick, consisting of parenchyma and minor veins (parenchyma enriched region). The slices were collected in a buffer solution consisting of 5 mM 2-(N-morpholino)ethansulfonic acid (MES)/Tris (pH6.0), 20 mM KCl, 0.5 mM $CaCl_2$, 0.2% (w/v) bovine serum albumin (BSA) and 0.2%(w/v) polyvinylpyrrolidone (PVP). The osmolality of the solution was adjusted to 150 $mOsmolkg^{-1}$ with sorbitol. The slices were washed at 4^{o}C to remove cell debris by two 3-min washes in 27 ml of the above buffer solution (about three slices per ml). The slices were divided into three groups (each group contained 27 slices) and pretreated with 27 ml buffer solution with or without 10 mM N-ethylmaleimide (NEM), 0.5 mM Erythrosine B (EB) (in 2003) or 300 mM KNO_3 (in 2004) in shaking water bath at 25^{o}C for 10 min. Excess NEM and EB were removed by three 3-min washes with 20 ml of the buffer solution at 4^{o}C. Each set of NEM and EB (or KNO_3) treated slices was then gently blotted dry on three layers of tissue paper and divided into two replicates. Each replicate was transferred to 5 ml of buffer solution containing 1 mM [^{14}C] glucose ($1.85x10^{-3}$ $MBq.ml^{-1}$, specific activity $1.85x10^{-3}$ $GBq.mol^{-1}$, Sigma Aldrich, Saint Louis, Missouri) and shaken for 10 min at 25 ^{o}C. The ^{14}C-label in the free space tissue was removed by three 3-min washes at

4 °C. Tissue fresh weights were recorded. The soluble ^{14}C-activity accumulated by the mesocarp tissues was extracted with 10 ml of 80% (v/v) aqueous ethanol overnight at 80 °C till dryness. The ethanol-soluble fractions were then radioassayed by adding 0.5 ml of water and 4.5 ml scintillation fluid (6 g of 2,5-diphenyloxazole, 500 ml of Triton X-100 and 1 L toluene). Radioactivity was determined using a Beckman LS6500 Liquid Scintillation Analyzer (Beckman Instrument, CA, U.S.A.) with automatic quench correction. For the radioassay of the ethanol-insoluble fraction, ethanol extracted slices were digested with NCS solubilizer (Amersham International, Amersham, Bucks., UK) at 50 °C for 48 h. The ratio of tissue per solubilizer was one to six (w/v). After this incubation 0.5 ml of water and 4.5 ml of scintillation fluid were added and the radioactivity was detected as described previously. The radioactivity was expressed as μmol mg protein $^{-1}$ h^{-1}.

4.3 Results

4.3.1 Fruit growth

'Sunset' papaya fruit weight was steady during the period 90 to 139 days after anthesis (DAA) in 2003 and 90 to160 DAA in 2004 (Figure 4.1), whereas 'UH801' fruit increased continuously in weight and reached maximum weight at 153 DAA in 2003 and 160 DAA in 2004. Fruit from 'UH810' was approximately 3 times larger than 'Sunset' at the same age. However, fruit weight of 'UH801' was less between 118 and 132 DAA during the dry weather of the warm season in 2003.

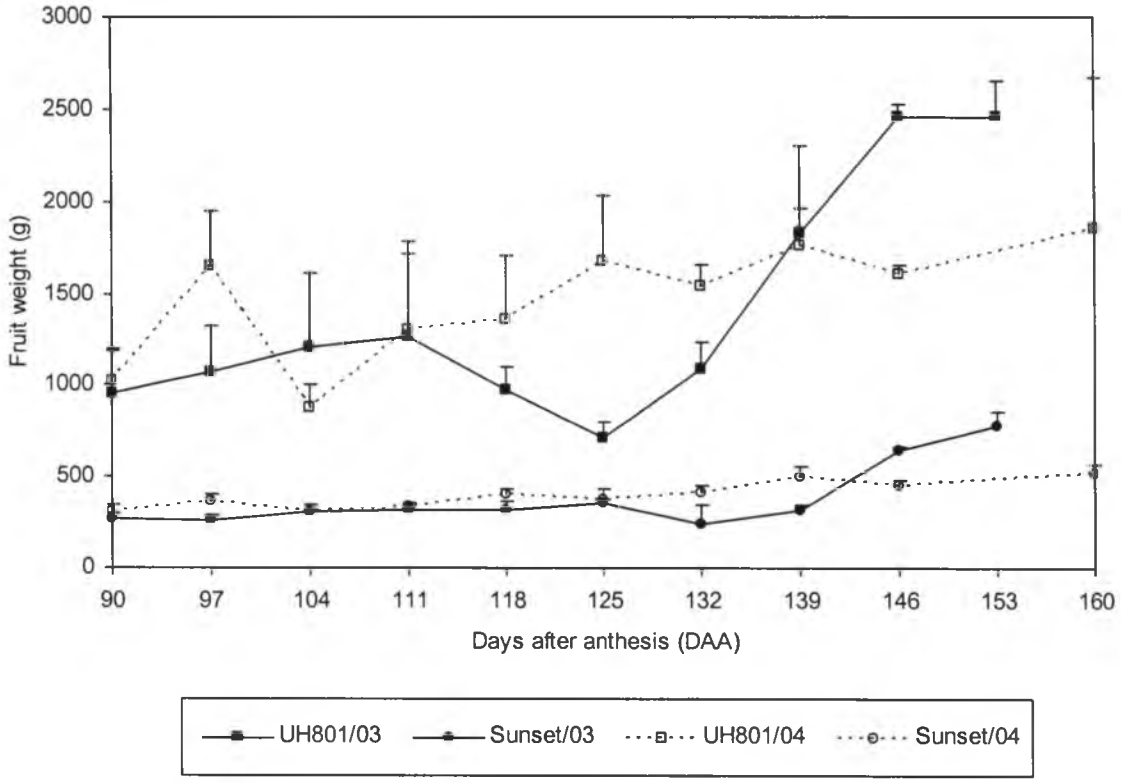

Figure 4.1. The fruit weight of cv Sunset and UH801 between 90 to 160 DAA harvested in 2003 and 2004. Mean ± SD, n = 3.

4.3.2 Color development

Skin color change started to occur in both 'Sunset' and 'UH801' at the same time in both years, 125 DAA (2003) and 139 DAA (2004) (Figure 4.2). The percentage of yellow skin on 'Sunset' at 160 DAA in 2004 was double than of 'UH801'.

Both 'Sunset' and 'UH801' have red orange pulp. Flesh color of 'Sunset' started to develop 111 DAA (2003) (Figure 4.3), 7 to 14 days before 'UH801', but both occurred at 125 DAA in 2004, 14 days later than in 2003. In both years, papaya flesh color started to develop 14 days ahead of skin color break (Figure 4.2).

The color of the seed sacrotesta changed from white to gray to brown and black. Similar to flesh color development, seed color of 'Sunset' started to develop 104 DAA (2003) (Figure 4.4), 7 days earlier than 'UH801', but at the same time (139 DAA) in 2004 (Figure 4.4).

4.3.3 Fruit firmness

Flesh firmness of both cultivars did not change between 97 to 132 DAA in 2003 (Figure 4.5). 'Sunset' fruit started to soften from 132 DAA, one week before 'UH801' in 2003. However, fruit firmness of both cultivars when harvested in 2004 was lower during 111 to 132 DAA and after 146 DAA periods.

4.3.4 Dry weight accumulation

'Sunset' flesh dry mass between 90 to 153 DAA harvesting in 2003 was greater than 'UH801'. Dry mass of 'Sunset' increased 111 DAA (6.8%), then remained constant between 118 to132 DAA (8.4-9.3%) before increasing again and reached a maximum 146 DAA (14.2%) (Figure 4.6). The dry weight of 'UH801' increased slightly after 132 DAA (6.0%) and was three weeks slower than 'Sunset' fruit in increasing in dry weight. In 2004, dry weight accumulation of both cultivars began to increase at the same time (139 DAA). The pattern of dry weight accumulation for both cultivars was similar to TSS accumulation in both seasons (Figure 4.7).

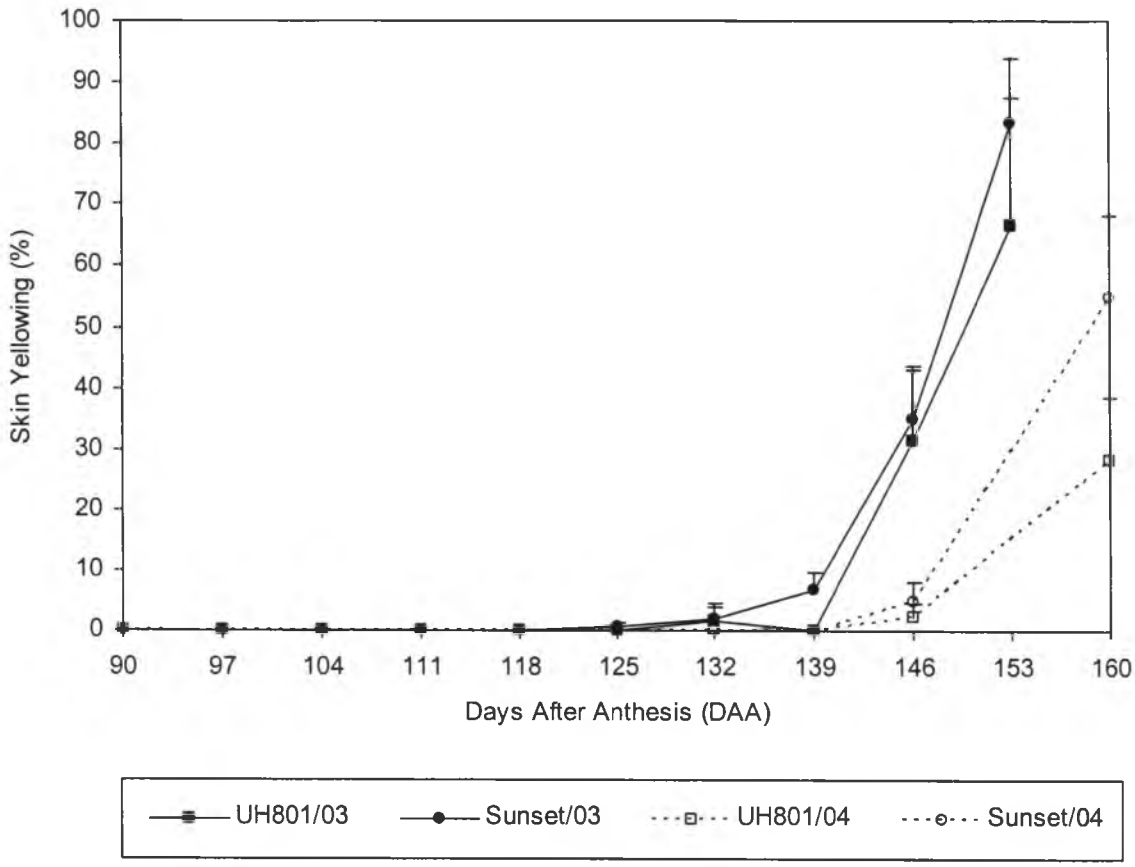

Figure 4.2. Skin color development of cv Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

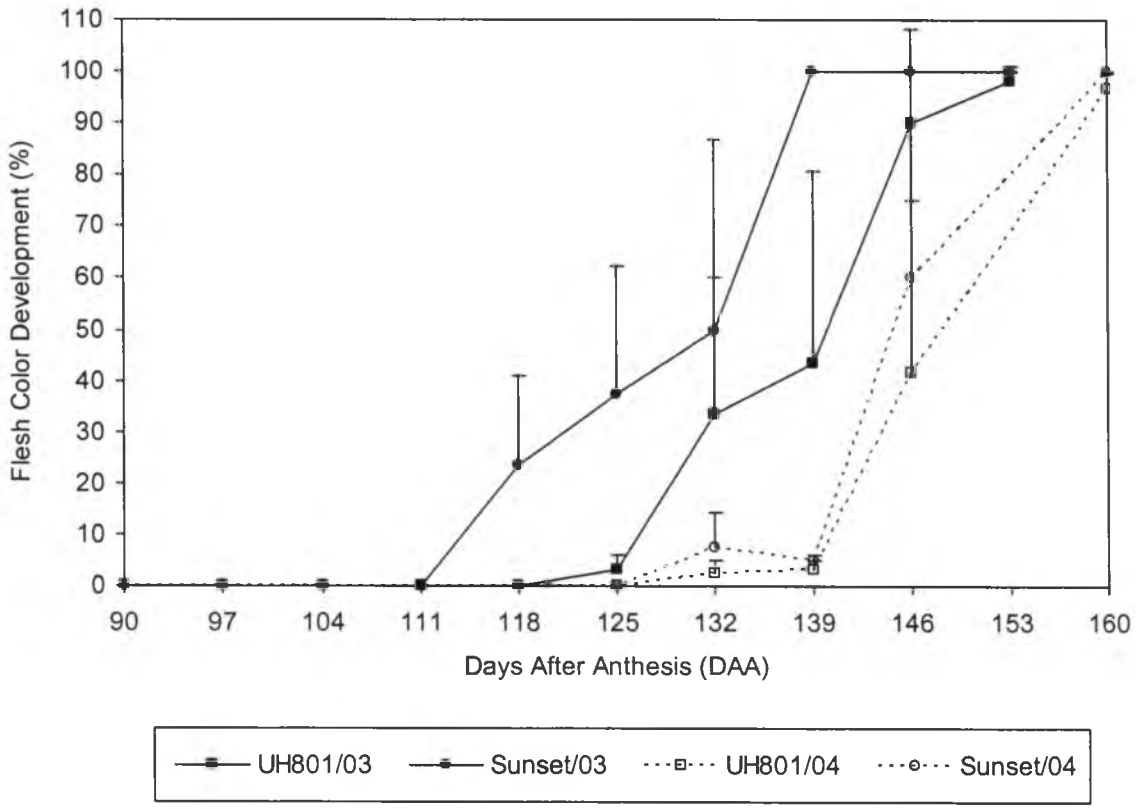

Figure 4.3. Flesh color development of cv Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

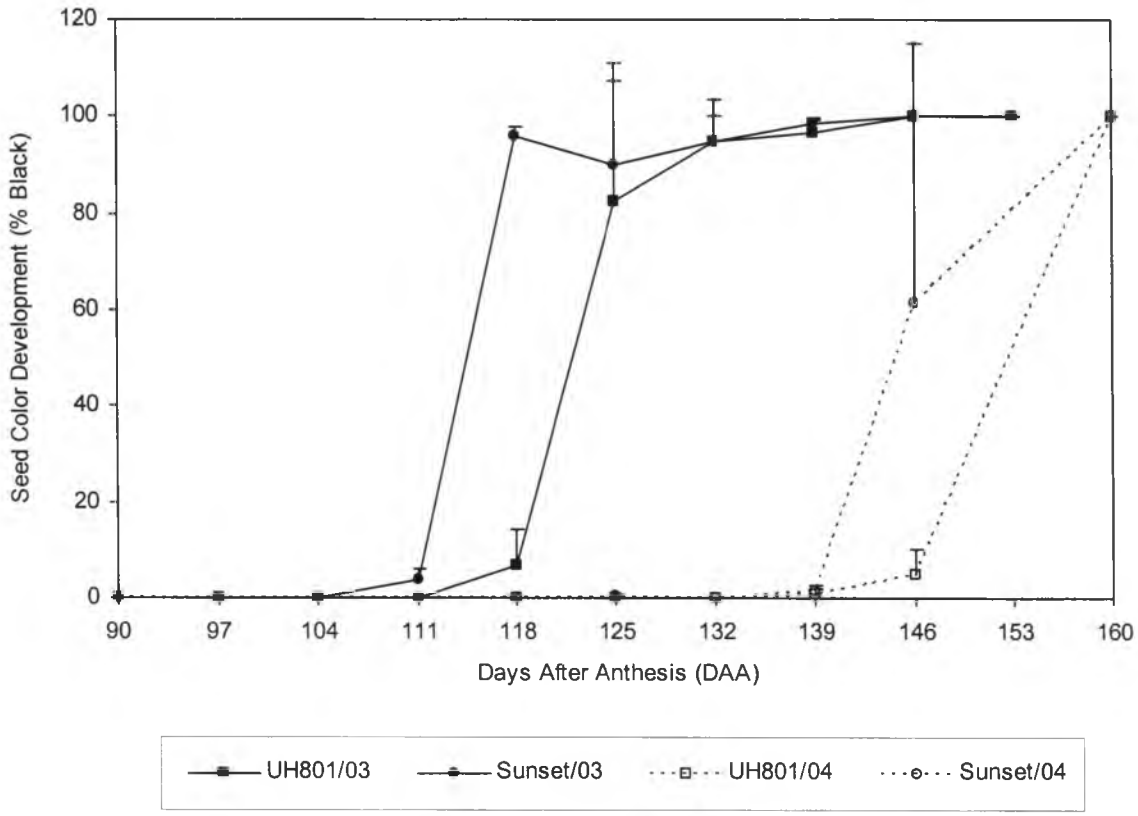

Figure 4.4. Seed color development of cv Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

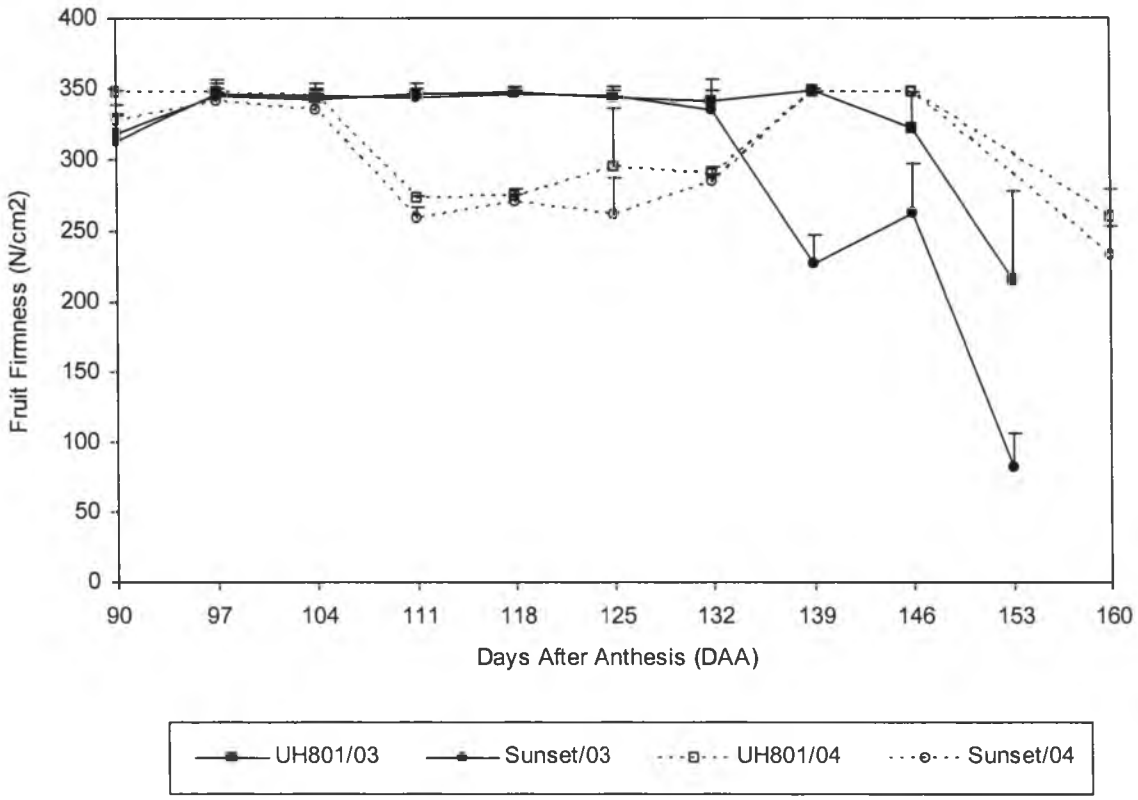

Figure 4.5. Fruit firmness of cv Sunset and UH801 harvested in 2003 and 2004. Mean ± SD, n = 3.

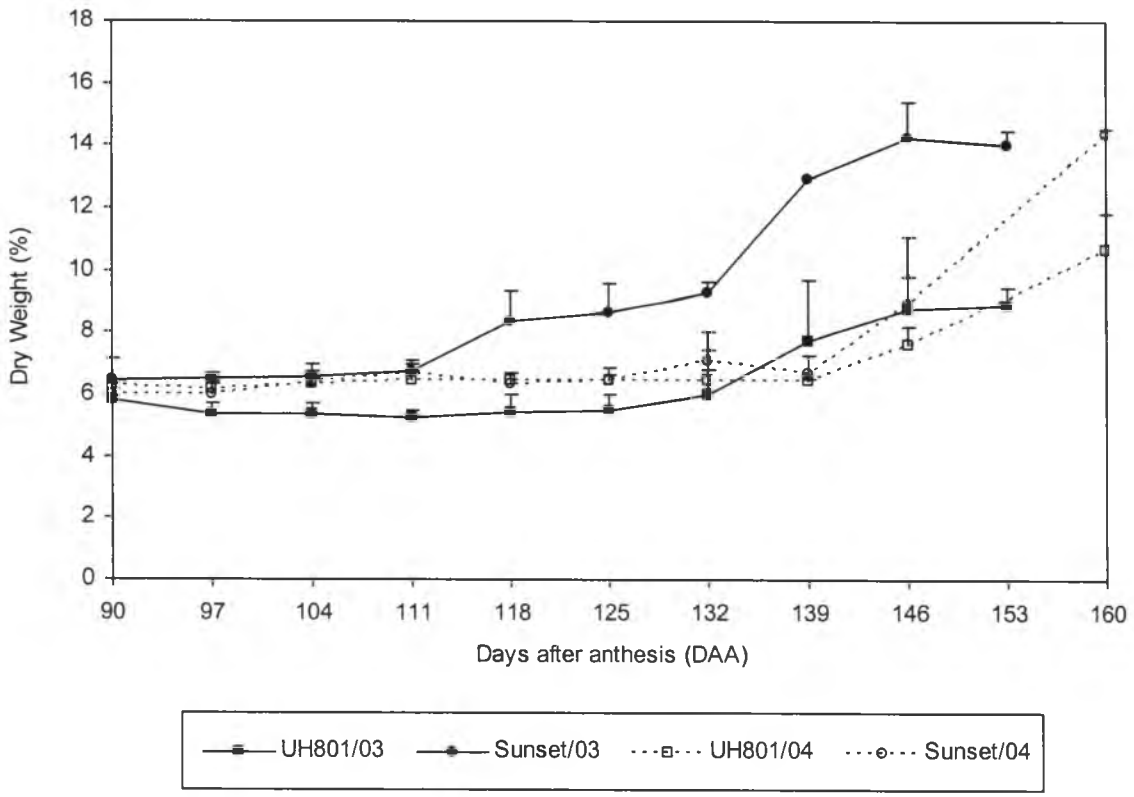

Figure 4.6. Dry weight accumulation of cv Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

4.3.5 Total soluble solids accumulation

Total soluble solids (TSS) (Figure 4.7) began to increase at the same time as flesh color development (Figure 4.3) and dry weight accumulation (Figure 4.6). At 111 DAA TSS was 5.5% in 'Sunset' and at 132 DAA was 5.3% in 'UH801', in 2003, and at the same time 139 DAA in 2004. TSS of 'Sunset' papaya fruit harvested in 2003 remained constant from 118 to 132 DAA (8.4-8.9%) before rapidly increasing again and reached maximum at 153 DAA (15.9%). The TSS of 'UH801' papaya fruit increased slightly three weeks after 'Sunset' fruit, however, TSS of both cultivars harvested in 2004 increased at the same date, 139 DAA.

4.3.6 Total sugar content

Total sugar content of both cultivars grown in 2003 dropped from 17 and 19 mg gFW^{-1} at 90 DAA to 4 and 7 mg gFW^{-1} at 97 DAA in 'UH801' and 'Sunset', respectively, and then gradually increased from 97 until they reached a maximum at 153 DAA (Figure 4.8). TSS was constant between 90 to 139 DAA in fruit grown in 2004 before rapidly increasing after 139 DAA.

4.3.7 Total protein accumulation

Total protein of both cultivars grown in 2003 increased slightly between 90 to 118 DAA (Figure 4.9). However, the protein content rapidly increased after 118 DAA in 'Sunset' fruit, whereas in 'UH801' it remained low until 139 DAA before rapidly increasing during the 146 to 153 DAA period. The increase in total protein content of both cultivars grown in 2004 occurred at the same date (139 DAA) and reached the maximum at 160 DAA.

4.3.8 ^{14}C-glucose uptake (μmol mg protein^{-1}h^{-1})

Hexose transporter activity was detected in both varieties but ^{14}C-glucose uptake in both the ethanol insoluble fraction and the total uptake did not appear to correlate with papaya sugar accumulation. In 2003, total glucose uptake by 'Sunset' fruit discs was lower than for 'UH801' (Figure 4.10), but it was higher in 2004 (Figure 4.11). The pattern of both the total ^{14}C-glucose uptake and the uptake in the ethanol insoluble fraction by 'Sunset' papaya mesocarp

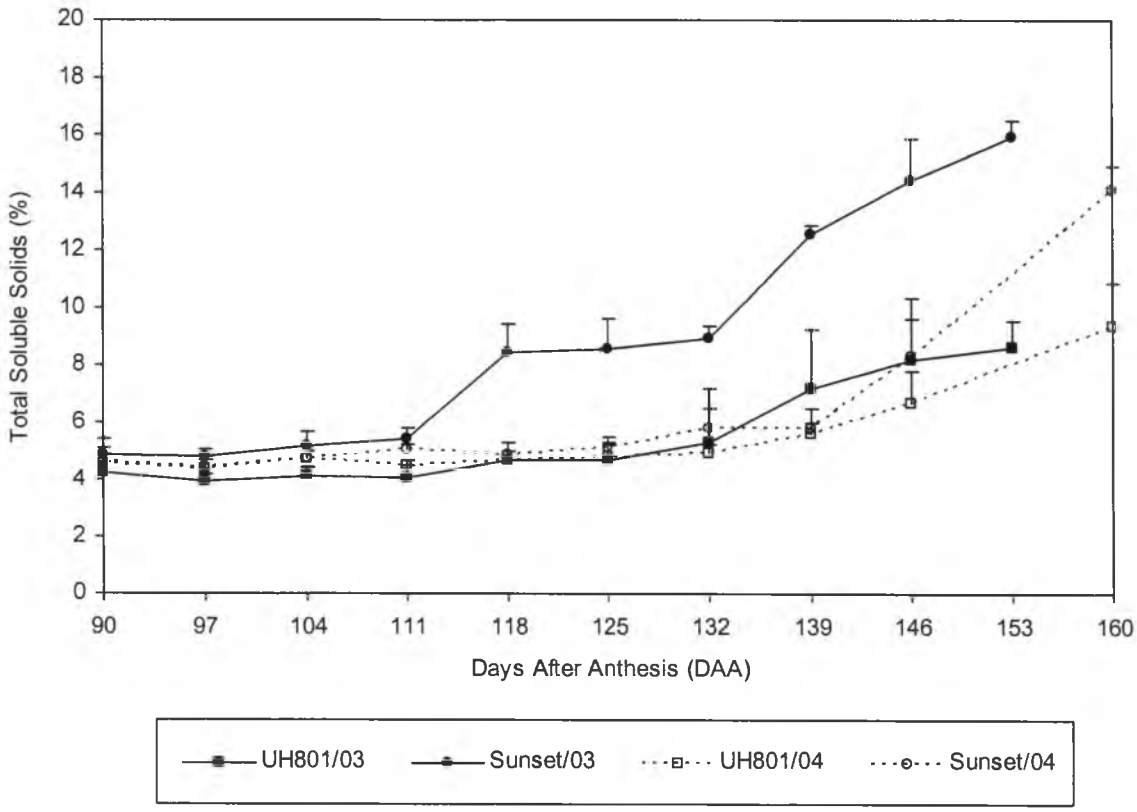

Figure 4.7. Total soluble solids (TSS) of cv Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

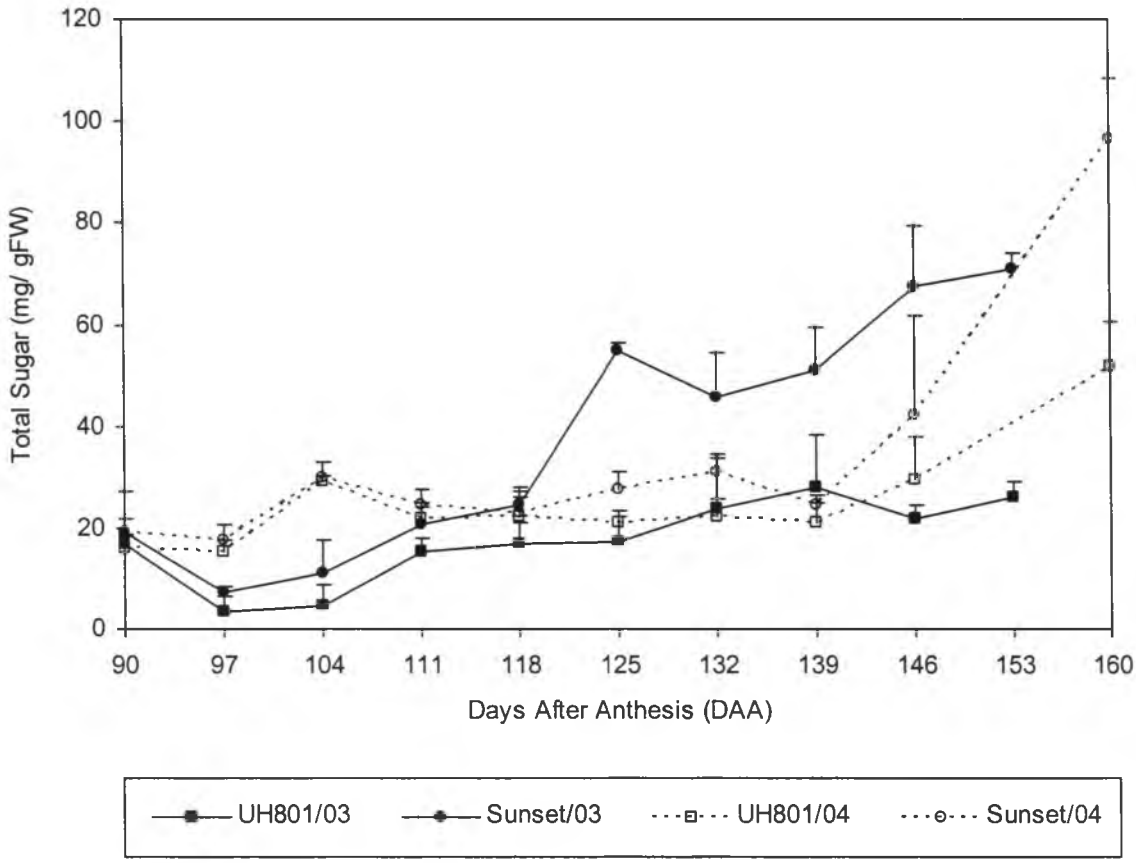

Figure 4.8. Total sugar content of cv Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

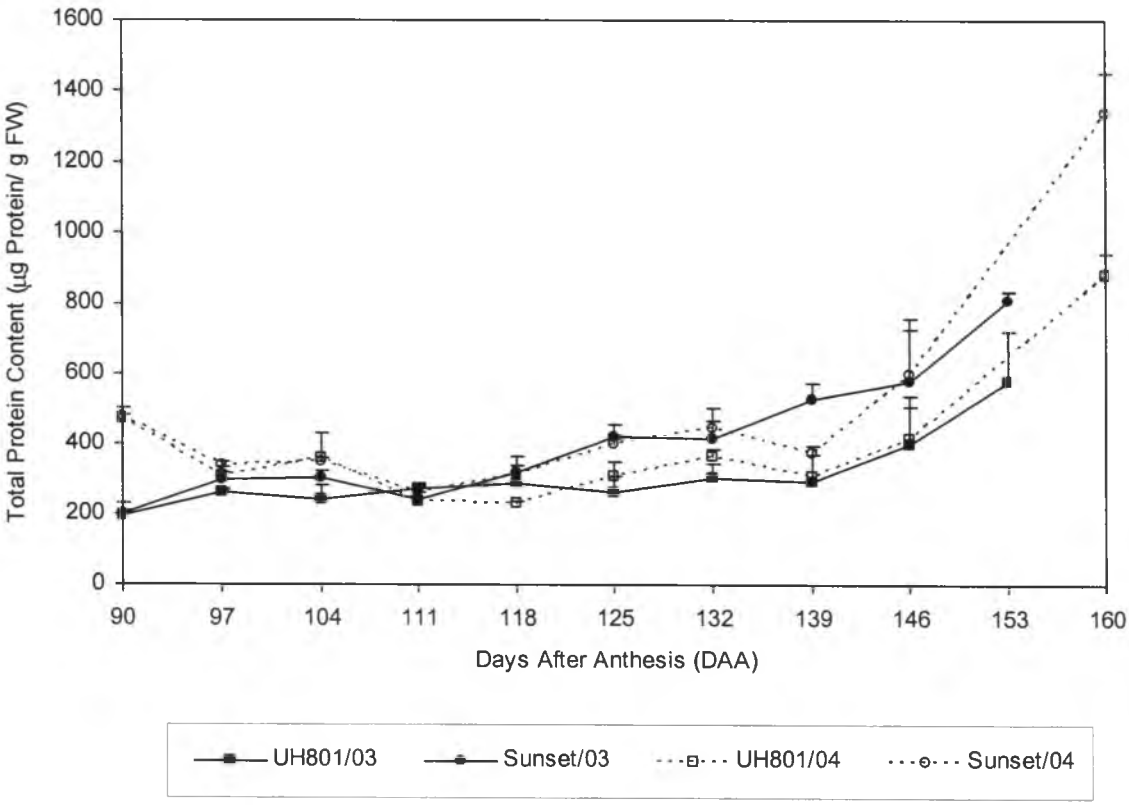

Figure 4.9. Total protein content of cv. Sunset and UH801 fruit harvested in 2003 and 2004. Mean ± SD, n = 3.

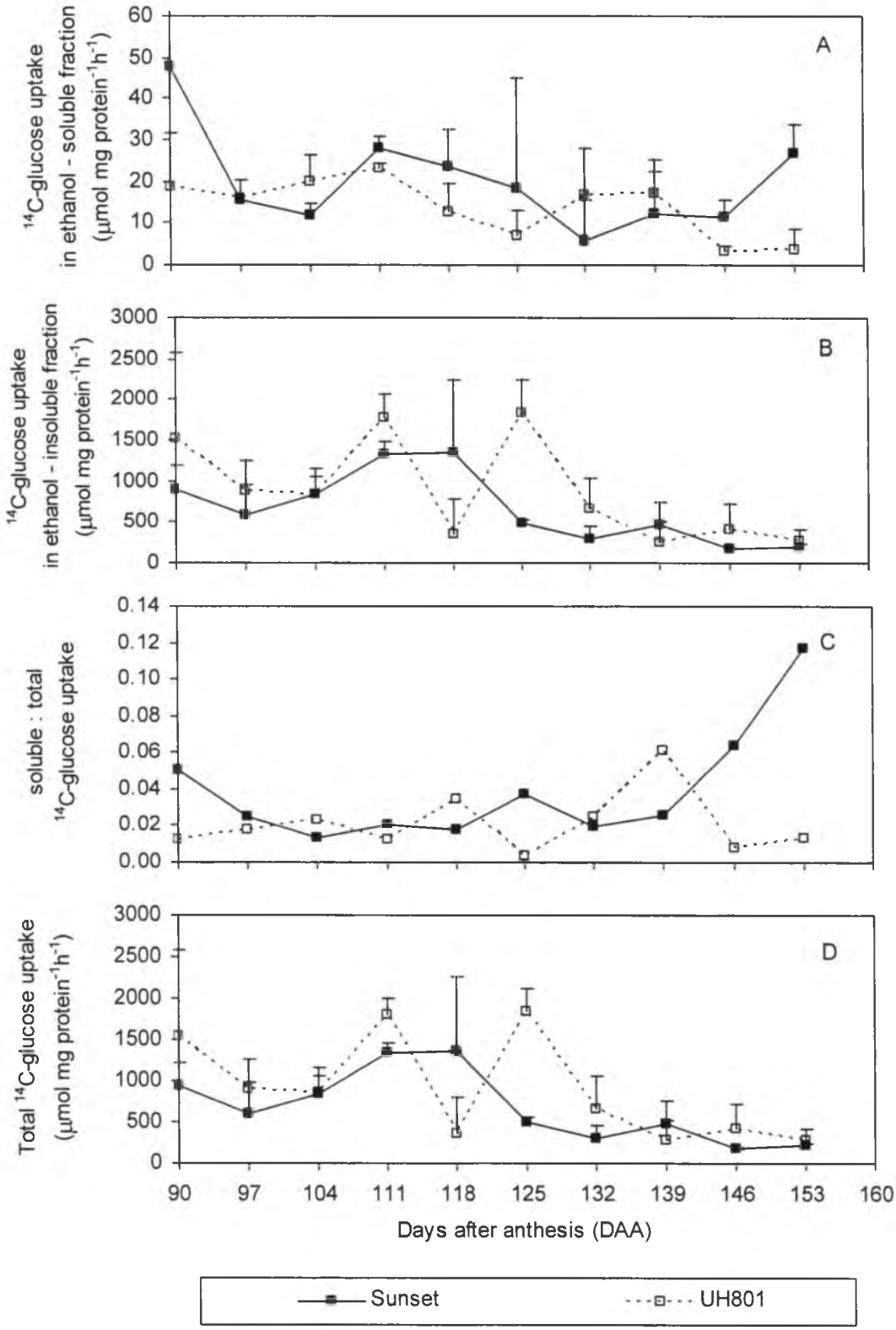

Figure 4.10. ^{14}C-glucose uptakes in sunset papaya mesocarp harvested in 2003 A) ethanol-soluble fraction, B) ethanol-insoluble fraction, C) soluble: total uptake ratio and D) total uptake.

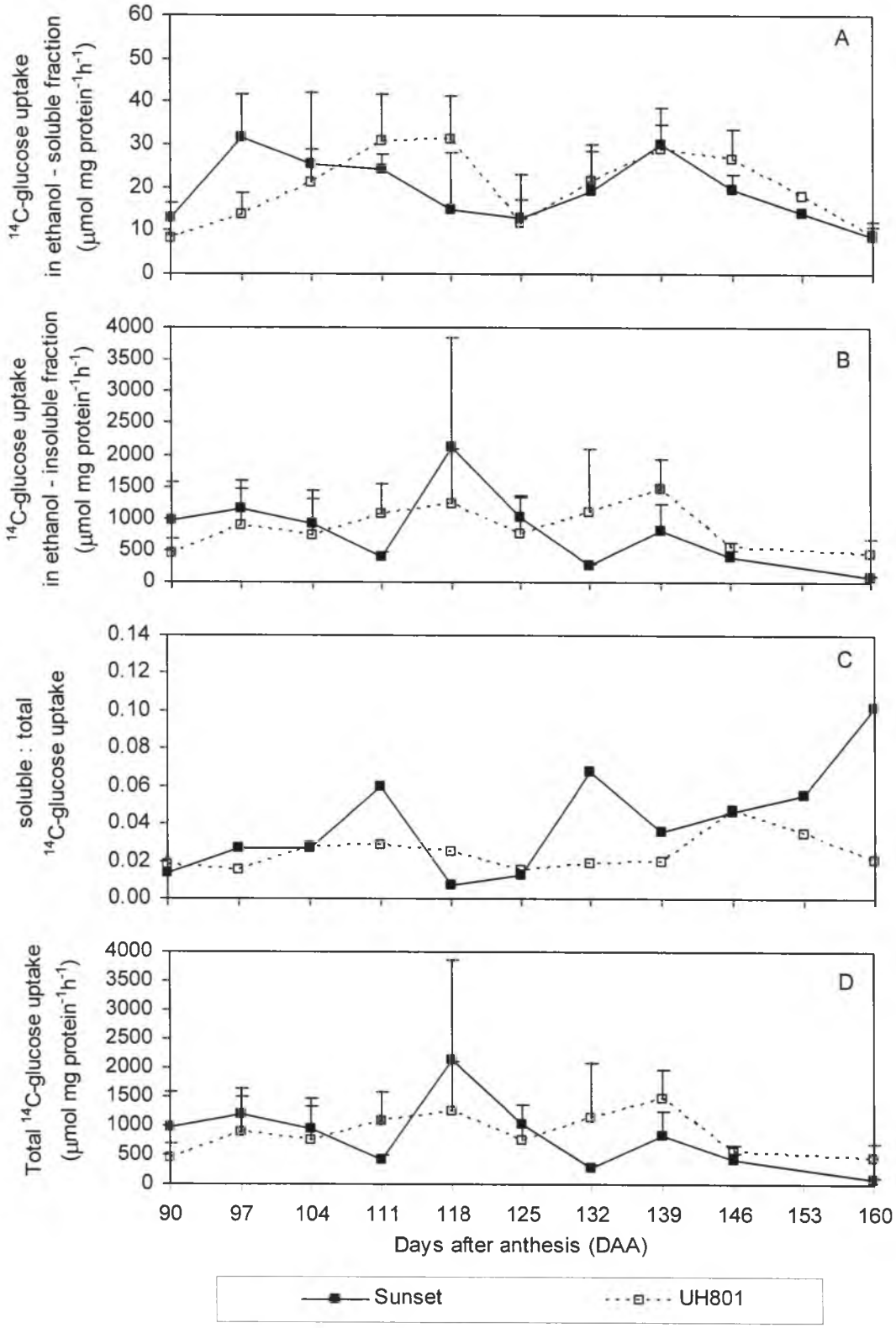

Figure 4.11. ^{14}C-glucose uptakes in sunset papaya mesocarp harvested in 2004 A) ethanol-solubie fraction, B) ethanol-insoluble fraction, C) soluble: total uptake ratio and D) total uptake.

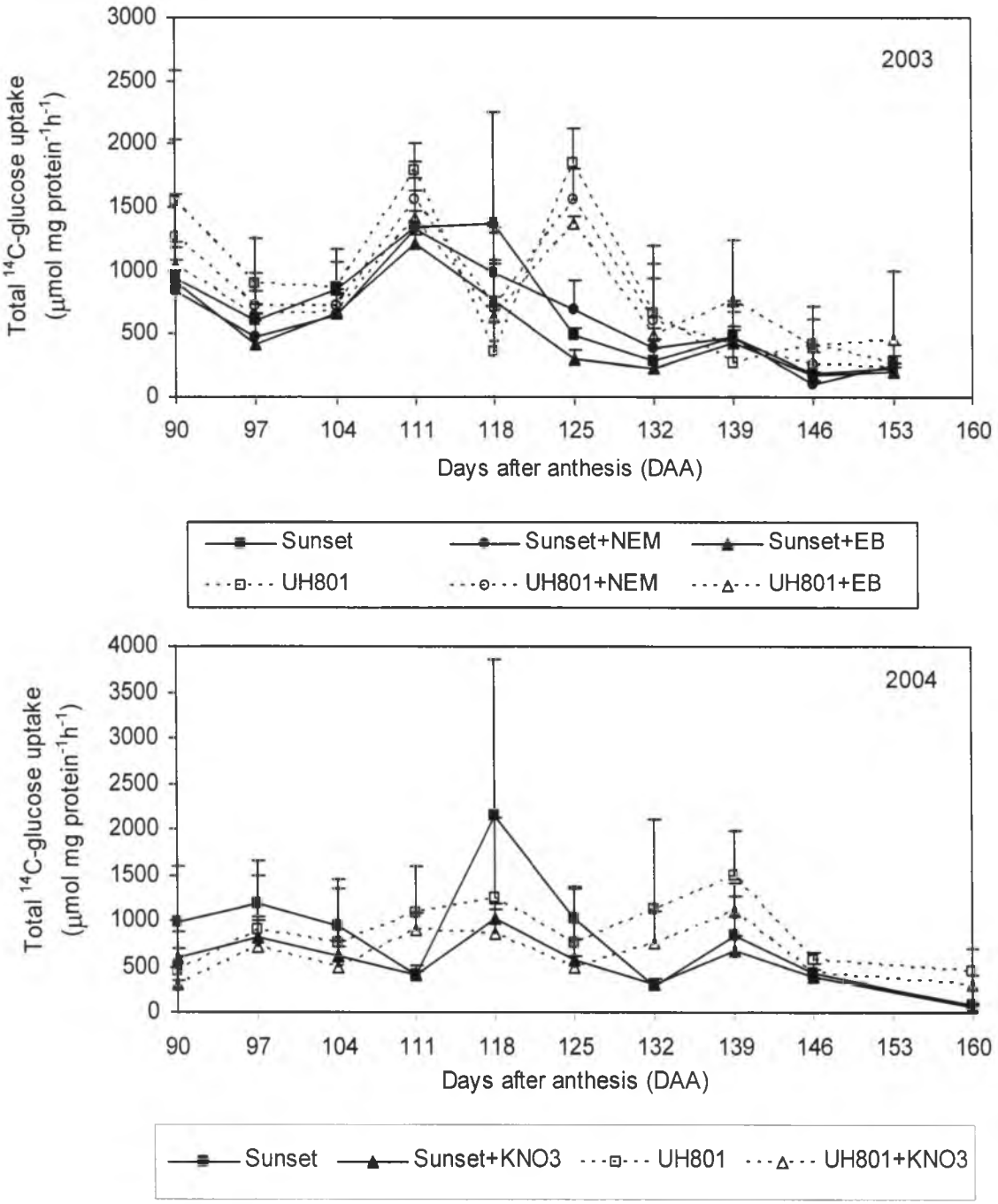

Figure 4.12. The total uptake of ^{14}C-glucose (µmol mg protein^{-1}h^{-1}) by 'Sunset' and 'UH801' papaya fruit discs harvested in 2003 and 2004 with and without inhibitors.

Table 4.1. Total C^{14}-glucose uptake (µmol mg protein^{-1}h^{-1}) of papaya flesh discs cv Sunset and UH801 between 90 to 153 DAA during 2003.

Age	90		Average	Pr	97		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	940	1547	1243		609	894	751	
10mM NEM	833	1266	1050		478	718	598	
0.5mM EB	902	1042	972		422	646	534	
Average (cv.) [1/]	891	1285		0.2138	503 b	753 a		0.0429
Pr (Treatment)			0.7533				0.2924	
Pr (Cul x Trt)		0.8156				0.9731		

Age	104		Average	Pr	111		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801		
Control	850	862	856		1346	1793	1569	
10mM NEM	643	723	683		1324	1553	1438	
0.5mM EB	661	677	669		1216	1419	1317	
Average (cv.) [1/]	718	754		0.7242	1295 b	1588 a		0.0237
Pr (Treatment)			0.2655				0.2317	
Pr (Cul x Trt)		0.9516				0.6369		

Age	118		Average	Pr	125		Average [1/]	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	1367	360	863		497	1846	1171 a	
10mM NEM	975	706	841		693	1560	1127 a	
0.5mM EB	765	641	703		303	1375	839 b	
Average (cv.) [1/]	1036	569		0.0930	498 b	1594 a		<.0001
Pr (Treatment)			0.8593				0.0159	
Pr (Cul x Trt)		0.3516				0.1096		

Age	132		Average	Pr	139		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	294	670	455		481	274	392	
10mM NEM	394	604	484		480	423	456	
0.5mM EB	228	488	340		429	765	573	
Average (cv.) [1/]	305	587		0.0740	463	488		0.8515
Pr (Treatment)			0.7011				0.3807	
Pr (Cul x Trt)		0.8948				0.2289		

Age	146		Average	Pr	153		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	181	412	296		227	279	253	
10mM NEM	105	258	182		255	247	251	
0.5mM EB	169	403	286		196	467	332	
Average (cv.) [1/]	152b	358a		0.0200	226	331		0.3449
Pr (Treatment)			0.4288				0.7855	
Pr (Cul x Trt)		0.8897				0.5491		

[1/] Data were analyzed by Duncan's multiple range test, means followed by the same letter were not significantly different.

Table 4.2. Total C^{14}-glucose uptake (μmol mg protein^{-1}h^{-1}) of papaya flesh discs cv Sunset and UH801 between 90 to 160 DAA during 2004.

Age	90		Average	Pr	97		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	975	446	711		1193	903	1048	
300mM KNO_3	599	299	449		815	718	767	
Average (cv.) [1/]	787	373		0.0807	1004	811		0.4431
Pr (Treatment)			0.2434				0.2743	
Pr (Cul x Trt)		0.5946				0.6960		

Age	104		Average	Pr	111		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	
Control	946	750	848		410	1090	750	
300mM KNO_3	619	488	554		404	900	652	
Average (cv.) [1/]	783	619		0.5330	407b	995a		0.0054
Pr (Treatment)			0.2751				0.5442	
Pr (Cul x Trt)		0.8998				0.5698		

Age	118		Average	Pr	125		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	2140	1250	1695		1022	763	893	
300mM KNO_3	1026	865	946		574	490	532	
Average (cv.) [1/]	1583	1058		0.3760	798	627		0.4423
Pr (Treatment)			0.2182				0.1274	
Pr (Cul x Trt)		0.5340				0.6914		

Age	132		Average	Pr	139		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	280	1128	704		843	1482	1163	
300mM KNO_3	304	763	533		679	1099	889	
Average (cv.) [1/]	945	292		0.0593	761b	1291a		0.0374
Pr (Treatment)			0.5822				0.2339	
Pr (Cul x Trt)		0.5308				0.6216		

Age	146		Average	Pr	160		Average	Pr
	Sunset	UH801	(treatment)	(cultivar)	Sunset	UH801	(treatment)	(cultivar)
Control	424	569	496		85	445	265	
300mM KNO_3	394	428	411		71	303	187	
Average (cv.) [1/]	409	498		0.1394	78b	374a		0.0052
Pr (Treatment)			0.1555				0.3447	
Pr (Cul x Trt)		0.3413				0.4361		

[1/] Data were analyzed by Duncan's multiple range test, means followed by the same letter were not significantly different.

discs were divided into three fruit growth phases during fruit maturation. Uptake was initially low during the first, 90-97 DAA (2003) and 90-111 DAA (2004) periods, then high 97-132 DAA and 111-132 DAA in 2003 and 2004, respectively. In the last phase, after 132 DAA for both years, the uptake was low. The maximum total hexose uptake in 'Sunset' was at 118 DAA in both years, 1,367 and 2,140 μmol mg protein^{-1}h^{-1}, respectively, in 2003 and 2004. Two peaks occurred in 'UH801' in 2003; 111 DAA (1,793 μmol mg protein^{-1}h^{-1}) and 125 DAA (1,846 μmol mg protein $^{-1}$ h^{-1}). In 2004, the peaks occurred at 118 DAA (1,250 μmol mg protein $^{-1}$ h^{-1}) and 139 DAA (1,482 μmol mg protein $^{-1}$h^{-1}). However, the glucose uptake ratio between soluble to total uptake in 'Sunset' papaya increased during 139 to 160 DAA in both years while it decreased in 'UH801'.

Erythrosin B (EB) (Figure 4.12, 2003) and KNO_3 (2004), inhibitors of plasma membrane and tonoplast H^+-ATPase, appeared to inhibit glucose uptake when glucose uptake rate was high, however, it was ineffective during the low uptake phase. During the low uptake phase, KNO_3 seemed to have a greater effect on uptake rate than EB.

4.4 Discussion

The growth pattern of papaya fruit has been characterized as a double sigmoid curve (Muda, *et al.*, 1994; Qui *et al.*, 1995; Zhou, 1999). The color break stage of both 'Sunset' and 'UH801' papaya fruit started 125 DAA (2003) and 139 DAA (2004), two weeks later than fruit grown in summer of 2003 (Figure 4.2) as found by Zhou (1999). Zhou (1999) reported that 'Sunset' papaya fruit grown during the warm season matured 140 DAA and the most obvious changes were skin, flesh and seed colors. The flesh color of 'Sunset' started to develop at 111 DAA (Figure 4.3), 7 to 14 days earlier than 'UH801'. 'Sunset' flesh color was fully developed at 139 DAA and it took 153 DAA for 'UH801' in 2003, and 160 DAA in 2004 for both cultivars. The warm weather in 2003 during fruit growth accelerated papaya fruit skin and flesh color development similar to that reported by Nakasone (1986).

The pattern of dry weight (DW) and total soluble solid (TSS) accumulation were similar for both cultivars and was similar to the results of Zhou (1999). Since papaya does not

store starch (Chen *et al.*, 2001), the dry matter content in the papaya flesh is mainly TSS. There are three major principal sugar forms in papaya flesh: glucose, fructose, and sucrose (Sankat and Maharaj, 1997). In the early stage of papaya fruit development, glucose is the major sugar form whereas the sucrose content is low but it becomes the major sugar form in mature and ripe fruit, about 80% of the total sugar (Chan *et al.*, 1979). Both the dry weight and TSS accumulation of 'Sunset' papaya was greater than for 'UH801'. In 2003, both dry mass and TSS contents increased at 111 DAA, was constant between 118 to132 DAA before increasing again to reach a maximum 146 to 153 DAA (Figure 4.6 and 4.7). The dry weight of 'UH801' slightly increased after 132 DAA (6.0%) and was three weeks slower than 'Sunset' fruit. However, dry weight and TSS contents in both cultivars increased at the same time in 2004 after 139 DAA. This result was possibly due to the slow growth rate during the cool season of 2004. Nakasone (1986) reported that cool weather extended the growth period of papaya growing in Hawaii by about two weeks. Both cultivars stop growing almost or at the same time. The unloaded sugar that was once metabolized for cell growth and development during fruit growth was now accumulated in the vacuole (Genard *et al.*, 2007; Yamaki and Asakura, 1988). 'Sunset' papaya fruit growing in the warm season of 2003 reached maturity and earlier sugar accumulation about two weeks before 'UH801'.

Total sugar content of the two cultivars dropped from 17 and 19 mg gFW^{-1} at 90 DAA to 4 and 7 mg gFW^{-1} at 97 DAA in 'UH801' and 'Sunset' in 2003, respectively. Total sugar then increased continuously to reach the maximum at 153 DAA (Figure 4.8), however, TSS was constant during 90-139 DAA in fruit growing in 2004 before rapidly increasing after 139 DAA similar to flesh and seed color development (Figure 4.3 and 4.4), dry weight and TSS accumulation (Figure 4.6 and 4.7), and total protein accumulation (Figure 4.9). The difference in total sugar accumulation rates between the two seasons was possibly due to the temperature as cool weather is known to delay fruit development (Nakasone, 1986; Nakasone and Paull, 1998). Zhou and Pauli (2001) reported that sugar accumulation in 'Sunset' papaya fruit occurred after seed maturation and during late fruit development, (100-140 DAA) and was related to an increase in acid invertase (AI) activity. AI activity began to increase at 90 DAA before sugar accumulation

had started and had the highest activity at 125 DAA. Sucrose phosphate synthase (SPS) and sucrose synthase (SS) activities, the other two main enzymes involved less sucrose metabolism, remained iow during this late maturation stage (Zhou and Paull, 2001).

Hexose transporter activity in papaya mesocarp tissue was studied with slices agitated in ^{14}C-glucose solution. The result showed that the pattern of total ^{14}C-glucose uptake in papaya had peaks in activity (Figure 4.10 and 4.11) so that the pattern of uptake could be divided into three phases. Total glucose uptake was low 90-97 DAA (2003) and 90-111 DAA (2004) in the first phase and during the last phase, after 132 DAA for both years. The peak in activity occurred between 97-132 DAA (2003) and 111-132 DAA (2004). The maximum total glucose uptake in 'Sunset' was found at 118 DAA in both years at 1,367 and 2,140 μmol mg protein^{-1}h^{-1}, respectively. Two peaks occurred in 'UH801' in 2003; 111 DAA (1,793 μmol mg protein^{-1}h^{-1}) and 125 DAA (1,846 μmol mg protein $^{-1}$ h^{-1}) just prior to the increase in soluble solids (Figure 4.7) and total sugar content (Figure 4.8). In 2004, 118 DAA the uptake rate was 1,250 μmol mg protein^{-1}h^{-1} and at 139 DAA 1,482 μmol mg protein^{-1}h^{-1} prior to and during the increase in TSS and total sugar content. The increase in 'Sunset' glucose uptake (Figure 4.10 and 4.11) paralleled the increase of acid invertase activity reported by Zhou and Pauli (2001). The peak in 'Sunset' hexose transporter activity occurred one week prior to the increase in acid invertase activity (125 DAA). This increase in total uptake was due to the increase in glucose uptake in ethanol insoluble fraction which indicated that transported glucose was used in cell wall formation during 97-132 DAA and 111-132 DAA in 2003 and 2004, respectively. The high soluble to total uptake ratio after 139 DAA (Figure 4.10 C and 4.11 C) indicated that glucose uptake during fruit ripening probably was not being used for structural syntheses but was being stored as soluble form in the fruit mesocarp and this correlated with papaya sugar accumulation.

Total glucose uptake by 'UH801' declined from 1,793 to 360 μmol mg protein $^{-1}$ h^{-1} between 111 DAA and 118 DAA in 2003. This decline could have been due to an error of tagging or harvesting as new papaya flowers occur every ± 2-3 days and usually 3 to 6 flowers open each week on an individual plant. Fruit at age 111 and 118 DAA could therefore have overlapped. The

physical characteristics of 'UH801' (skin and flesh color and firmness) and chemical compositions (dry weight, TSS, total sugar content and total protein) between these two stages were similar and only the glucose uptake was dramatically different. A second explanation would be that since the fruit were sampled during dry season, some fruit and some trees had died before the experiment could be finished. The effect of dry weather can be seen on fruit growth (Figure 4.1). Sample fruit were collected from the remaining trees; fruit at 118 DAA were collected three weeks after those at 125 DAA. Therefore, the activity of glucose transporter in papaya fruit at 118 DAA might be less than fruit at 125 DAA (Figure 4.10).

Although, the total glucose uptake in 'UH801' fruit slices was higher than for 'Sunset' (Figure 4.10), TSS (Figure 4.7) and total sugar content (Figure 4.8) were less than in 'Sunset'. Moreover, Zhou and Pauli (2001) reported that the activity of acid invertase of 'UH801' was significantly lower than 'Kopoho', a high sugar variety, during late fruit development. The results for 'Sunset' had an overall lower hexose transporter activity than 'UH801' and therefore hexose transporter was possibly not the main factor controlling sugar accumulation in papaya fruit. However, 'Sunset' fruit size and weight (Figure 4.1) were much smaller than 'UH801' and 'Sunset' fruit stopped growing and started to mature earlier than 'UH801'. It was possible that 'Sunset' had a smaller cell size and volume and/or number of cell per fruit than those of 'UH801', a low sugar breeding line. Therefore, 'Sunset' fruit tissue would possibly have a lower need for sugar to produce energy for cell division and enlargement during fruit growth and also for a cell wall formation which occurred after cell division and before enlargement has been completed (Taiz and Zeiger, 2002). The sugar remaining after completing fruit growth would be expected to move via hexose transporter and stored in the vacuole. The relationship between fruit quality, such as fruit size and composition, and the complex chain of biological processes including tissue differentiation and cell functioning (cell division, endoreduplication, expansion, metabolic transformations, and vacuole storage) has been recently discussed (Genard *et al.*, 2007). The same relationship has also found in the uptake of sorbitol and other sugars in the apple fruit flesh (Yamaki and Asakura, 1988).

Transferring hexose sugar across the plasma-membrane and the tonoplast of

storage parenchyma is an active process in matured tomato fruit (Ruan and Patrick, 1995). Erythrosin B (EB) inhibits plasma-membrane H^+-ATPase activity in many plant tissues (Cocucci, 1986; Beffagna and Romani, 1988). EB was effective in inhibiting uptake when glucose uptake was high 104-132 DAA (Figure 4.12) (Table 4.1). In contrast, there was little or no effect on glucose uptake during the phase of low hexose accumulation 90-97 and 139-153 DAA. The result suggests that papaya hexose transporter is an energy-coupled reaction. The inhibitory effect of KNO_3, a reversible inhibitor of tonoplast (V-type) H^+-ATPase, on glucose uptake in 2004 confirmed the 2003 results of the inhibitory effect of EB. In addition, KNO_3 seems to have greater effect on hexose uptake than EB.

The tomato hexose transporter (*LeHT2*) is strongly inhibited by the membrane – permeable SH-group modifier *N*-ethylmaleimide (NEM) (Gear *et al.*, 2000). Ma *et al.* (2002) also found that NEM inhibited proton pumping activity of tonoplast H^+-ATPase of *Populus euphratica,* a halophytic plant. However, in this study, *N*-ethylmaleimide (NEM) was found to be less effective in inhibiting uptake (Figure 9). The overall inhibition pattern of NEM was similar to that of EB inhibition. The mode of action of NEM is completely different from *p*-chloromercuriphenylsulfonic acid (PCMBS), a membrane-impermeable SH-group modifier (M'Batchi and Delrot, 1984; Gear *et al.*, 2000). Hexose transporter inhibition by NEM may be effective on the inner surface of plasma-membrane or tonoplast or may be due to the indirect effect on energy metabolism rather than direct effect on the transporter (Xia and Saglio, 1988; Gear *et al.*, 2000). In papaya fruit flesh, hexose uptake inhibition was not sensitive to NEM. Papaya plasma-membrane and tonoplast hexose transporters may not contain sulhydryl group on the cytoplasmic side of the transporter. Unlike NEM, PCMBS inhibits hexose uptake at the outer surface of plasma-membrane. PCMBS binds to the active site of the transporter (M'Batchi and Delrot, 1984). PCMBS is highly effective in inhibiting hexose uptake of tomato during late fruit development (Ruan and Patrick, 1995). However, due to its high toxicity PCMBS is no longer available. More research on the optimum condition, inhibition, and the effect of a proton gradient on hexose transporter activity in papaya, and molecular studies of this protein are needed.

4.5 Summary

'Sunset' papaya fruit had smaller fruit size but contained higher total soluble solids, total sugar, dry weight and total protein concentration than those of UH801, a low sugar content breeding line. Maturation and sugar accumulation of 'Sunset' occurred two to three weeks prior to accumulation in UH801. Papaya hexose transporter appeared to be energy-dependent. The activity of hexose transporter of 'Sunset' fruit was lower than those of UH801 in 2003, but was higher in 2004. Hexose transporter was probably not the main factor controlling papaya sugar accumulation.

CHAPTER 5

MOLECULAR CHARACTERISTIC OF HEXOSE TRANSPORTER GENE OF PAPAYA FRUIT

5.1 Introduction

Hexose transporters have been investigated at the molecular level in numerous plants (Ruan *et al.*, 1997; Shiratake *et al.*, 1997; Weber *et al.*, 1997; Fillion *et al..*, 1999; Gear *et al.*, 2000). Plant hexose transporters belong to a large multigene superfamily of transmembrane facilitators (MFS, major facilitator superfamily) (Marger and Saier, 1993) consisting of at least two members in peach fruit (Etienne *et al.*, 2002), eight members in castor bean (*Ricinus communis*) (Weig *et al.*, 1994) and 14 genes in *Arabidopsis thaliana* (BÜttner *et al.*, 2000; Williams *et al.*, 2000).

Each transporter gene is expressed in different tissues and organs at different developmental stages. In grape berry cv. Pinot Noir, two hexose transporters, *VvHT1* and *VvHT2,* have been cloned (Fillion *et al.*, 1999). *VvHT1* is expressed in the berries and in young leaves. Expression of *VvHT1* in the berries is highest shortly after anthesis, then declines, and increases again after the ve´raison stage, the stage of the onset of maturity when the berries change color and soften. In white varieties (like Chardonnay), ve´raison is identified by softening and berry translucency. In red varieties, ve´raison is seen when the berries turn purple (Fillion *et al.*, 1999). However, the second peak of *VvHT1* expression does not always occur (Terrier *et al.*, 2005). *VvHT2* is expressed in the berries when *VvHT1* is expressed at a low level. The expression of *VvHT2* starts about 2 weeks before ve´raison stage and peaks about 1 week after this stage. However, the level of *VvHT2* expression is low compared with *VvHT1* (Fillion *et al.*, 1999). In tomato, both *LeHT1* and *LeHT3* are highly express in young fruit and root tips while *LeHT2* shows a high expression in source leaves and flowers (Gear *et al.*, 2000). Some genes, such as the *Arabidopsis AtSTP4* (Truernit *et al.*, 1996) are expressed during plant stress.

Hexose transporters have been described from both the plasma membrane and tonoplast (Rausch, 1991; Shiratake *et al.*, 1997). The hexose transporters have 12

transmembrane helices (Marger and Saier, 1993). In tomato, the full length *LeHT2* cDNA encodes a protein of 523 amino acids, with a calculated molecular mass of 57.6 kDa. The predicted tomato protein has 12 putative membrane-spanning domains similar to the other members found in the MFS (Gear *et al.*, 2000).

In papaya, Zhou and Paull (2001) reported the increase in acid invertase (AI) activity during papaya fruit maturation and proposed that sugar unloading in papaya fruit during maturation is possibly by an apoplastic pathway. The mechanism of sugar unloading via the apoplast pathway requires the presence of hexose transporter(s), however, there is no evidence that these transporters occur in papaya fruit during papaya sugar accumulation. The absence of this data means that the importance of apoplastic unloading during sugar accumulation and the role of invertase is unsupported.

The objective of current study was to characterize hexose transporter present during fruit ripening. This characterization involved isolating a full length cDNA clone and determining the polypeptide sequences of papaya hexose transporter. The papaya sequence was compared to the hexose transporter gene and the peptide with other plants. The presence of transmembrane regions and their location, and the structure of this protein were predicted. The final aspect was to determine the expression of the hexose transporter gene during fruit maturation and ripening of the commercial cultivar Sunset.

5.2 Materials and methods

5.2.1 Papaya tissues

Three papaya fruit ('Sunset') were harvested weekly after the seeds began to mature, about 90 days after anthesis (DAA) (±3 days), until the full ripe stage about 153 DAA during August to November 2006. Tissue samples were collected from the fruit mesocarp, frozen in liquid nitrogen and stored at -80°C until used.

5.2.2 Fruit characterization

Papaya fruit weight of cv Sunset was measured weekly and reported in gram per fruit (g $fruit^{-1}$). The dry mass percentage of each sample was determined by drying ten gram of fruit flesh tissue at 60-70^{O}C for at least 72 hrs or until the weight was stable.

The skin color development was subjectively determined by visual estimation of the area of skin yellowing (0-100% yellow). The flesh color development was subjectively estimated as the degree of internal flesh showing carotenoid development (0-100% orange). Seed color development was subjectively estimated from the blackening and sacrotesta formation (0-100% black). The percentage of the total soluble solid (TSS) was measured by a digital handheld refractometer (Palm Abbe Model AR200, MISCO, Cleveland, Ohio).

5.2.3 RNA isolation

Total RNA was isolated according to the method of Mason and Botella (1997) from frozen fruit pericarp tissue at different stages from 90 DAA to 160 DAA. Ten grams of frozen fruit tissue were ground to a powder in liquid nitrogen with a coffee grinder and transferred to a weighted centrifuge bottle. Hot phenol (2.5 volumes) was immediately added and the mixture was shaken for one min at room temperature. The extraction buffer (3 volumes) (150 mM Tris base, 2 % (w/v) SDS, 50mM EDTA, 1 % (v/v) β-mercaptoethanol, adjusted pH to 7.5 with boric acid) was added and the tube was shaken again for one min. Pre-chilled absolute ethanol (0.25 volumes) and 5M potassium acetate (0.11 volumes) were added and shaken for one min at room temperature. One volume of chloroform-isoamyl alcohol (24:1) was added and the final mixture was shaken for another min. The homogenate was then centrifuged at 27,300 g (13,000 rpm with Sorvall GSA rotor) at 4°C for 40 min. The upper phase was transferred to a new bottle and extracted with water-saturated phenol (pH 4.5)-chloroform-isoamyl alcohol (25:24:1) by shaking at room temperature for one min and centrifuged for 20 min at the same speed. The upper phase was again transferred to a new bottle and the phenol-chloroform-isoamyl alcohol extraction was repeated. This extraction step was repeated two to three times or until no interphase layer was observed. Pre-chilled absolute ethanol (2.25 volumes) was added and then allowed to precipitate

at -20°C for two hours or overnight. The nucleic acids were collected by centrifugation at 27,300 g for 40 min at 4°C. The pellet was washed with pre-chilled 80% ethanol and again centrifuged at 27,300 g for 20 min. The pellet was air dried for 10 min, re-suspended in 7.5 ml RNase-free water (DEPC treated water), precipitated in a 50 ml-centrifuged tube with lithium chloride (a final concentration of 2M) and allowed to stand at -20°C overnight. The RNA was collected by centrifugation at 24,464 g (13,000 rpm with Sorvall SA-600 rotor) for 40 min at 4°C. The pellet was washed with pre-chilled 80% ethanol and centrifuged again for 20 min. The ethanol was completely removed. The pellet was air dried and re-suspended in 300 μl of RNase-free water. The suspension was transferred to a clean Eppendorf vial, mixed with pre-chilled absolute ethanol (3 volumes) and 3M sodium acetate (0.1 volumes) and incubated at -80°C for 30 min. The RNA solution was centrifuged at 15,700 g (13,000 rpm with Eppendorf 5415R) for 30 min at 4°C. The pellet was washed twice with pre-chilled 80% ethanol. The ethanol was completely removed and air dried. The RNA pellet was redissolved in 50 μl of RNase-free water.

The absorbance of the total RNA was measured at 260 nm (1 optical density = 40 μg/ml). The quality of the total RNA was determined by electrophoresis on formaldehyde gels. The remaining total RNA solution was directly used for RT-PCR, poly A^+ mRNA isolation and Northern blot analysis or stored in -80°C until use.

5.2.4 Reverse Transcription and PCR ampiifications (RT-PCR)

The first strand cDNA was synthesized from 10 μg of total RNA of the immature green 'Sunset' fruit tissue (118 DAA) using Oligo (dT) primer following the instruction of StrataScript™ Reverse Transcriptase (Stratagene, La Jolla, CA, Cat. No.600085). Degenerate primers were designed from conserved regions of plant hexose transporter cDNA sequences obtained from GenBank for barrel medic (*Medicago truncatula MtST1,* U38651), *Arabidopsis thaliana AtSTP1* (X55350), tobacco (*Nicotiana tabacum NtMST1,* X66856), fava bean (*Vicia faba VfSTP1,* Z93775), castor bean (*Ricinus communis RcSCP,* L08196), grape berry (*Vitis vinifera VvHT,* Y09590), and tomato fruit (*Lycopersicon esculentum LeHT1,* AJ132223, and *LeHT3,*

AJ132225) (Figure 5.1).

The degenerate sequence of hexose transporter primers were as follow: forward primer (HxTF486) 5'- GGH TAC GAT ATH GGR ATY TC - 3' (Tm 49°C); and reverse primer (HxTR1076) 5'-TBV CCM CGY TCR ATC ATK GA - 3' (Tm 57°C). Papaya gene specific primers for acid invertase were used as positive control. The sequences of the invertase primers were as follow: forward primer (PpyInvF234) 5'- TGG ATT AAC GAC CCA AAT GC - 3' (Tm 53°C) and reverse primer (PpyInvR787) 5'- AAT CTG GGC ATT CCC ACA TA - 3' (Tm 53°C). The amplification system was as follow:

Component	Volume (µl)
Sterile, deionized water	13.8
10x PCR buffer	2.5
10 mM dNTP mix	0.5
Taq DNA polymerase (Promega)	0.2
25 mM $MgCl_2$	2.0
cDNA template	4.0
25 µM Forward primer	1.0
25 µM Reverse primer	1.0
Total volume	25.0

The thermocycling regime was as follows: step 1 initial denaturation at 95°C for 5 min; step 2 denatured at 95°C for 1 min; step 3 annealed at gradient temperature between 45-56°C for 1 min; step 4 extension at 72°C for 1 min, PCR at steps 2 to 4 was repeated for 35 cycles, followed by final extension period (step 5) of 72°C for 7 min using the Eppendorf Mastercycler® personal 5332. The PCR products were 553 and 590 base pairs (bp) in size, for invertase and hexose transporter, respectively and were purified from a 1% agarose (1xTAE) gel using QIAEX II Gel Extraction Kit (QIAGEN).

5.2.5 Cloning of the hexose transporter cDNA fragment

The DNA fragments obtained from PCR were ligated into the pGEM®-T Easy vector (Promega, Madison, Wisconsin) at the ratio of 3:1 for insert DNA: vector. The cloning reaction was performed following the manufacturer procedures and incubated at 16°C overnight. The ligated DNA/vector (5 µl) was transformed into 50 µl One Shot® TOP 10 competent cells (Invitrogen, Carlsbad, California) following the manufacturer procedure. The transformed colonies were selected by blue/white screening from the LB (Luria-Bertani) medium plus 100 µg/ml ampicillin and 40 µl 2% X-Gal agar plate after incubation at 37°C overnight. The selected-transformed colonies were then inoculated into 2 ml LB liquid medium plus 100 µg/ml ampicillin and agitated at 37°C overnight. The transformed competent cells were transferred into Eppendorf vials and precipitated at 15,700 g (13,000 rpm with Eppendorf 5415R) at room temperature for 5 min. The plasmid DNA was then purified using QIAprep Spin Miniprep Kit (QIAGEN, Valencia, California). The amount of plasmid DNA was detected by reading the absorbance at 260 nm. The quality of the plasmid DNA was determined by cutting with Not1 restriction enzyme (NEBcutter, New England BioLabs) overnight at 37°C and then running on 1% agarose/ 0.5xTBE gel electrophoresis.

5.2.6 Sequence analysis of the hexose transporter gene

Plasmid DNA was sequenced by the 96-capillary 3730xl DNA Analyzer (Applied Biosystems, Hitachi, Japan) at the Center for Genomics, Proteomics, and Bioinformatics Research Initiative, University of Hawaii at Manoa. The first papaya hexose transporter sequence was searched in BLASTn against the NR (all GenBank, no longer "non-redundant") database of NCBI (National Center for Biotecnology Information) to confirm that the PCR products have homology to known monosaccharide transporters.

Figure 5.1 Multiple alignment of hexose transporter genes from barrel medic (*Medicago truncatula MtST1,* U38651), *Arabidopsis thaliana AtSTP1* (X55350), tobacco (*Nicotiana tabacum NtMST1,* X66856), fava bean (*Vicia faba VfSTP1,* Z93775), castor bean (*Ricinus communis RcSCP,* L08196), grape berry (*Vitis vinifera VvHT,* Y09590), and tomato fruit (*Lycopersicon esculentum LeHT1,* AJ132223, and *LeHT3,* AJ132225). The conserved regions are underlined.

```
          1                                                    50
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    ATTTCTACAG TTTACGATTA CAGTTCTCTC TCAACTTTTT TCTTTTTTCT
MtST1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
VfSTP1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
RcSTC     ~~~~~~~~~~ ~CTCTGTTTA AGCTTCTTGT TCTATTTCTG TTTCTCTGCA
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    ~~~~~~~~~~ ~~~~~~~~~A AAAACCCATC CCATCAAAAA TAAACAAGAG

          51                                                   100
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    TGATCATTTC TTGTGCCACA AGTTGTTCTT GTTTTGTGAT CAGCTCAGAA
MtST1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~CTTCTT CTTGCTACTG AGGTCAGAAA
VfSTP1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~CTTCA TCGGAGAAAA
RcSTC     AGGAAGTGTA CGTAGCATAT AACTAAGAGC CAAAAAAGAA AAAAAAATGC
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    GGCCTAAAGA AGAATCCTAA AGACTTTACG GGTCTTGTTT AGGATAAAAG

          101                                                  150
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    AGATGGCTGG TGGTGGTGGT ATTGGTCCCG GCAACGGGAA AGAATATCCC
MtST1     AAATGGCTGG TGGTGGAATT CCCATTGGAG GGGGTAACAA AGAGTACCCC
VfSTP1    AAATGCCTGC AGCCGGAATC CCCATCGGAG CGGGGAACAA GGAGTACCCC
RcSTC     CTGCAGTAGG AGGTATACCG CCCTCTGGTG GCAACAGGAA AGTGTACCCG
VvHT      ~~ATGCCGGC TGTCGGAGGC TTTGATAAGG GTACCGGGAA GGCCTATCCC
AtSTP1    AAATGCCTGC CGGTGGATTC GTCGTCGGGG ATGGCCAAAA GGCTTATCCC

          151                                                  200
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    GGCAATTTAA CTCTTTATGT TACCGTTACG TGCATTGTCG CTGCCATGGG
MtST1     GGAAACCTCA CTCCTTTTGT CACCATAACA TGCATCGTTG CTGCCATGGG
VfSTP1    GGAAACTTAA CTCCTTTCGT CACCATAACA TGTGTTGTTG CAGCCATGGG
RcSTC     GGAAACCTTA CTCTTTATGT TACTGTAACA TGTGTCGTTG CAGCCATGGG
VvHT      GGTAACCTTA CTCCTTACGT GACTGTGACA TGTGTTGTTG CAGCCATGGG
AtSTP1    GGCAAACTCA CTCCCTTTGT TCTCTTCACT TGCGTTGTTG CTGCCATGGG
```

```
          201                                                       250
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    TGGTCTCATT TTCGGTTACG ATATTGGAAT TTCTGGAGGT GTGACATCAA
MtST1     TGGTTTGATC TTTGGCTACG ATATTGGAAT TTCAGGTGGT GTGACGTCCA
VfSTP1    TGGTTTGATC TTTGGTTACG ATATAGGAAT TTCAGGTGGT GTTACTTCAA
RcSTC     TGGCTTGATC TTTGGTTACG ATATTGGGAT TTCTGGTGGA GTTACGTCCA
VvHT      TGGTTTGATC TTTGGTTACG ATATTGGAAT TTCTGGTGGG GTCACGTCCA
AtSTP1    CGGTCTCATC TTCGGATACG ATATCGGAAT CTCCGGTGGT GTGACGTCTA
                     ~~~GGHTACG ATATHGGRAT YTC~~~
          251                                                       300
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    TGGACTCATT CTTGAGCAGA TTTTTCCCAT CTGTGTTCAG GAAGCAAAAG
MtST1     TGGATCCGTT TCTGAAGAAA TTTTTTCCGG CGGTGTACCG GAAAAAGAAC
VfSTP1    TGAATCCGTT TCTTGAGAAA TTTTTTCCGG CGGTGTACCG GAAGAAAAAC
RcSTC     TGGATTCATT CTTGAAGAAG TTCTTTCCTT CAGTTTACCG GAAGAAGAAA
VvHT      TGGCTCCGTT CTTGCAGAAG TTCTTCCCTT CTGTGTACCG GAAGGAGGCT
AtSTP1    TGCCGTCTTT CCTCAAGCGA TTCTTCCCGT CGGTGTATCG GAAACAACAA

          301                                                       350
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    GCAGATGATT CAACAAATCA ATACTGCAAA TTTGACAGCC AAACATTGAC
MtST1     AAGGACAAAT CGACAAACCA GTACTGTCAA TATGACAGTC AAACATTGAC
VfSTP1    GCGCAACATT CGAAGAATCA GTACTGTCAA TACGACAGTG AGACACTGAC
RcSTC     GCGGATGAAT CGTCAAACCA GTACTGTCAA TATGATAGTC AGACACTGAC
VvHT      TTGGACAAGT CCACGAATCA GTACTGTAAG TTTGATAGTG AGACACTAAC
AtSTP1    GAGGACGCGT CAACGAACCA GTACTGTCAG TACGATAGCC CGACGCTAAC

          351                                                       400
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    GATGTTCACG TCGTCATTGT ACTTGGCTGC TCTTTTGTCG TCTCTGGTGG
MtST1     GATGTTTACA TCGTCGTTGT ATCTGGCTGC CCTTTTGTCC TCGTTGGTAG
VfSTP1    CTTGTTCACA TCCTCGTTGT ACCTGGCCGC GCTTTTGTCG TCGGTGGTTG
RcSTC     TATGTTCACA TCTTCACTGT ACTTGGCTGC TTTAATTGCT TCTCTTGTAG
VvHT      GTTGTTCACA TCGTCGCTTT ATCTGGCTGC TCTTCTCTCG TCGCTGGTGG
AtSTP1    GATGTTCACA TCGTCTCTAT ATCTAGCGGC GCTAATTTCG TCGCTGGTGG

          401                                                       450
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    CATCTACTGT CACCAGAAAA CTTGGACGGA GACTTTCTAT GCTCTGTGGA
MtST1     CTTCCACCAT AACTCGTAGG TTTGGTCGGA AACTTTCCAT GCTTTTCGGA
VfSTP1    CTTCAACGAT CACTAGAAGG TTTGGTCGGA AACTCTCCAT GCTTTTCGGA
RcSTC     CATCTACTAT TACGAGAAAA TTTGGTAGGA AACTCTCTAT GCTTTTTGGC
VvHT      CCGCGACGGT GACCCGAAAG TTCGGGAGAA AGCTGTCAAT GCTATTCGGA
AtSTP1    CTTCCACCGT GACAAGAAAG TTCGGACGGC GGCTCTCGAT GCTCTTCGGC
```

Figure 5.1. (continue)

```
          451                                                        500
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    GGTGTCCTCT TCTGTGCTGG AGCTTTGATC AATGGCTTTG CTCAGAATGT
MtST1     GGATTACTTT TCCTTGTCGG TGCTCTTATT AATGGCTTTG CTAATCATGT
VfSTP1    GGCTTGCTTT TTCTGGTCGG TGCTCTCATT AATGGCCTTG CTCAAAACGT
RcSTC     GGTGTACTCT TTTGTGCTGG AGCTATCATC AATGGTGCGG CTAAAGCAGT
VvHT      GGACTGCTCT TTTGTGCCGG TGCCATCATC AATGGCGCTG CTAAAGCTGT
AtSTP1    GGCATACTCT TCTGCGCCGG AGCTCTCATC AATGGTTTCG CCAAACATGT

          501                                                        550
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    TGCTATGCTC ATTGTTGGTC GTATTTTACT AGGTTTTGGT ATTGGATTCG
MtST1     TTGGATGTTG ATCGTGGGTC GGATCTTGCT CGGGTTTGGT ATCGGGTTTG
VfSTP1    TGCGATGTTG ATCGTCGGTC GGATCTTGCT CGGATTCGGT ATCGGGTTTG
RcSTC     CTGGATGTTG ATTCTTGGTA GAATTTTGCT TGGTTTTGGT ATTGGGTTTG
VvHT      TTGGATGTTG ATTGTCGGTC GTATACTGCT GGGTTTTGGT ATTGGGTTTG
AtSTP1    TTGGATGCTC ATCGTCGGTC GTATCTTGCT TGGTTTCGGT ATCGGTTTCG

          551                                                        600
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    CCAATCAGTC TGTTCCACTA TACCTATCTG AAATGGCTCC ATACAAGTAC
MtST1     CTAATCAGCC TGTGCCATTG TACCTCTCTG AGATGGCTCC TTACAAGTAT
VfSTP1    CGAATCAGTC TGTGCCATTA TACTTGTCTG AGATGGCTCC ATACAAGTAC
RcSTC     CCAATCAGTC TGTGCCGCTC TACCTCTCTG AGATGGCTCC ATACAAATAC
VvHT      CCAATCAGTC TGTGCCGCTC TACCTCTCTG AGATGGCTCC ATACAAATAC
AtSTP1    CTAATCAGGC TGTGCCACTG TACCTCTCTG AGATGGCTCC ATACAAATAC

          601                                                        650
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    AGAGGAGCAC TCAACCTAGG TTTTCAACTG TCCATTACAA TTGGTATACT
MtST1     AGAGGAGCAT TGAATATTGG GTTTCAATTA TCAATTACAA TTGGTATACT
VfSTP1    AGAGGAGCGT TGAATATTGG ATTTCAATTG TCAATTACAA TTGGAATACT
RcSTC     AGAGGAGCAC TGAACATTGG TTTCCAGTTA TCAATTACAA TTGGTATCCT
VvHT      AGAGGAGCCC TCAACATTGG CTTCCAATTA TCCATCACAA TTGGTATTCT
AtSTP1    AGAGGAGCTT TAAACATTGG TTTCCAGCTC TCAATTACAA TCGGAATCCT

          651                                                        700
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    TGTAGCAAAT GTGTTGAACT ATTTCTTTGC CAAGATTCA. .....TTGGG
MtST1     TGTGGCCAAT GTGTTGAATT ACTTTTTTGC CAAAATCAAA GGTGGATGGG
VfSTP1    TGTGGCCAAT ATTTTGAACT ACTTTTTTGC CAAAATCAAA GGTGGATGGG
RcSTC     TGTAGCCAAT GTATTGAATT ACTTCTTTGC CAAGATTAAG GGTGGTTGGG
VvHT      TGTGGCCAAT ATATTGAACT ACTTCTTTGC AAAGATCAAG GGGGGTTGGG
AtSTP1    CGTCGCCGAA GTGCTAAACT ACTTCTTCGC CAAGATCAAA GGCGGTTGGG
```

Figure 5.1. (continue)

```
          701                                                       750
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    GATGGAGATT AAGCTTAGGA GGTGCTATGG TACCTGCATT GATCATCACA
MtST1     GATGGAGATT GAGTTTAGGT GGTGCTATGG TCCCAGCACT TATAATAACA
VfSTP1    GATGGAGATT GAGTTTAGGT GGTGCTATGG TCCCTGCACT TATAATAACC
RcSTC     GTTGGAGGCT GAGTCTTGGT GGTGCTATGG TCCCTGCCCT CATCATTACA
VvHT      GATGGAGATT GAGCTTGGGT GGCGCTGTGG TCCCTGCGCT CATCATCACC
AtSTP1    GATGGCGGCT CAGTCTCGGA GGCGCGGTGG TTCCTGCCTT GATCATAACC

          751                                                       800
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
NtMST1    ATAGGCTCAC TTTTCCTTCC CGAGACACCA AACTCCATGA TCGAACGTGG
MtST1     ATTGGATCAT TAGTCCTTCC CGACACCCCT AACTCAATGA TCGAACGTGG
VfSTP1    ATTGGTTCGC TAATCCTACC CGACACGCCA AATTCCATGA TCGAGCGTGG
RcSTC     GTTGGATCAT TAGTCCTTCC AGATACACCA AACTCCATGA TTGAACGGGG
VvHT      GTCGGGTCCC TTGTCCTCCC GGACACACCC AACTCCATGA TCGAGCGTGG
AtSTP1    ATCGGCTCCC TCGTCCTCCC TGACACTCCC AATTCAATGA TCGAGCGTGG
                                          3'~~AGKTACT ARCTYGCMCC

          801                                                       850
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ~AACCACGAC GAAGCCAAAG CTCGATTGAA GAGAATTAGG GGAATTGAAG
NtMST1    CAATCACGAC GAAGCCAAAG CTAGGCTTAA AAGAATCAGA GGCATTGATG
MtST1     TGATCGCGAT GGAGCTAAAG CTCAACTTAA GAGAATTCGC GGCATTGAAG
VfSTP1    GGATCGAGAT GGTGCCAAGG CACAGCTTAA GAGAATTCGC GGAGTTGAAG
RcSTC     CCAGCACGAA GAAGCCAGAG CACATCTAAA GAGAGTTCGC GGTGTAGAAG
VvHT      CCAGCACGAG GGAGCGAAAA CAAAACTGAG AAGAATCCGG GGTGTCGATG
AtSTP1    CCAACACGAA GAAGCCAAAA CCAAGCTCAG ACGAATCCGT GGTGTCGATG
          VBT~~5'

          851                                                       900
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     ATGTAGATGA AGAGTTCAAT GATTTGGTTA TTGCTAGTGA AGCTTCTAGG
NtMST1    ATGTAGACGA AGAGTTCAAT GATTTAGTCG TGGCGAGTGA GGCTTCTAGG
MtST1     ATGTTGATGA AGAGTTTAAT GACCTCGTAG CAGCTAGTGA GGCCTCAATG
VfSTP1    ATGTTGATGA AGAGTTTAAT GATCTTGTGG CTGCTAGTGA AACGTCGATG
RcSTC     ATGTTGATGA GGAGTTTACT GACCTTGTTC ATGCTAGTGA AGATTCAAAG
VvHT      ATGTTGAAGA GGAATTCAAT GACCTTGTTG TAGCCAGTGA GGCCTCCAAG
AtSTP1    ACGTCAGCCA AGAGTTTGAC GATTTGGTCG CCGCTAGTAA AGAGTCGCAG

              901                                                       950
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ AAAAGGAGAA ACAGGCCTCA
LeHT1     AAAATTGAAC ATCCCTGGAG GAACTTGTTG CAAAAGAAAT ATAGACCACA
NtMST1    AAAATTGAGA ACCCTTGGAG AAATTTGTTG CAAAGGAAAT ATAGGCCACA
MtST1     CAAGTTGAAA ACCCTTGGAG AAATTTGTTG CAGAGGAAAT ATAGACCTCA
VfSTP1    CAGGTTGAAA ATCCTTGGAG GAATTTGTTG CAGAGAAAAT ATAGACCTCA
RcSTC     AAAGTTGAAC ATCCTTGGAG GAATTTGTTA CAGAGGAAAT ACAGGCCTCA
VvHT      CTTGTTGAGC ACCCCTGGAG AAATCTCTTG CAGAGGAAGT ACAGGCCACA
AtSTP1    TCGATAGAGC ACCCGTGGAG AAACCTCCTC CGCCGCAAGT ACCGACCACA
```

Figure 5.1. (continue)

```
          951                                                      1000
LeHT3     GTTAATCATG GCGATAATGA TGCCGACTTT TCAGATACTT ACTGGCATTA
LeHT1     TCTTACAATG GCAATTATGA TCCCATTTTT CCAACAACTT ACTGGAATCA
NtMST1    TCTCACAATG GCAATTATGA TCCCATTTTT CCAGCAACTT ACTGGAATCA
MtST1     GCTTACTATG GCTGTATTGA TACCATTCTT CCAACAATTT ACAGGCATCA
VfSTP1    GCTTACTATG GCTGTGTTGA TTCCGTTCTT CCAACAGTTT ACTGGAATTA
RcSTC     TCTCTCAATG GCCATTGCAA TTCCGTTCTT TCAGCAACTC ACCGGCATTA
VvHT      CCTCACAATG GCCATCCTCA TTCCCTTCTT CCAGCAGCTT ACCGGGATTA
AtSTP1    TCTCACAATG GCCGTTATGA TTCCGTTCTT TCAACAGCTA ACCGGAATCA

          1001                                                     1050
LeHT3     ACATCATACT TTTTTATGCC CCAGTATTGT TTCAGAGTAT GGGGTTTAAA
LeHT1     ACGTGATTAT GTTTTATGCA CCTGTGTTGT TTAAAACCAT TGGTTTTGGT
NtMST1    ATGTGATTAT GTTCTATGCA CCAGTTTTGT TTAAGACTAT TGGTTTTGGT
MtST1     ATGTTATCAT GTTTTATGCA CCTGTGCTAT TTAATTCCAT TGGGTTTAAG
VfSTP1    ACGTTATTAT GTTTTACGCG CCTGTGTTGT TTAACTCGAT TGGGTTTAAG
RcSTC     ATGTGATCAT GTTCTATGCT CCTGTTTTGT TCGATACTAT TGGATTCGGT
VvHT      ATGTCATTAT GTTTTATGCC CCTGTTCTCT TCAAAACTAT TGGCTTTGCG
AtSTP1    ATGTGATTAT GTTTTACGCT CCGGTTTTGT TCAACACCAT TGGTTTCACG

          1051                                                     1100
LeHT3     AGAGCAGCCT CTCTGTATTC CTCTGCTTTG ACTGGTGCAG TTCTTGCTTC
LeHT1     ACTGATGCTT CACTTATGTC TGCTGTGATC ACTGGTGGAA TCAATGTCAT
NtMST1    GCTGATGCTT CCCTTATGTC TGCTGTTATT ACTGGTGGAG TCAATGTACT
MtST1     GACGATGCTT CACTTATGTC GGCTGTCATC ACCGGTGTTG TTAATGTTGT
VfSTP1    GATGATGCTT CACTTATGTC AGCTGTTATC ACCGGTGTTG TTAACGTTGT
RcSTC     AGTGATGCTG CACTCATGTC TGCTGTGATC ACTGGTCTTG TTAATGTTTT
VvHT      GATGATGCTT CCCTGATGTC TGCTGTGATA ACCGGGCGGG TTAATGTTCT
AtSTP1    ACCGATGCTT CTCTCATGTC CGCTGTGGTC ACTGGCTCGG TTAACGTTGG

          1101                                                     1150
LeHT3     ATCTACACTT TTATCAATGG CCACTGTCGA TAGATGGGGT CGAAGAGTTC
LeHT1     TGCCACTATT GTTTCTATTT ACTATGTTGA TAAATTAGGA AGAAGATTCT
NtMST1    TGCAACTGTT GTTTCTATTT ACTATGTTGA TAAATTGGGA AGAAGATTCT
MtST1     TGCTACTTGT GTCTCAATTT ATGGAGTTGA TAAGTGGGGT AGGAGAGCCC
VfSTP1    TGCTACTTGT GTCTCAATTT ATGGAGTTGA TAAGTGGGGG AGAAGAGCTC
RcSTC     TGCAACAATG GTCTCAATTT ATGGTGTTGA TAAGTGGGGA AGGAGGTTCC
VvHT      TGCAACCATA GTTTCAATCT ACGGTGTTGA TAAGTGGGTA AGAAGGTTTC
AtSTP1    CGCTACGCTT GTTTCTATCT ACGGTGTTGA CAGATGGGGA CGTCGGTTTC

          1151                                                     1200
LeHT3     TTCTTATTAC CGGTGGAATC CAAATGATCA TCTGTCAGGT TATTGTTGCG
LeHT1     TGTTTCTTGA AGGTGGAATT CAAATGCTCT TTTCCCAAAT AGCCGTGGCA
NtMST1    TGTTCCTTGA AGGTGGCATT CAAATGCTCA TCTGCCAAAT AGCGGTGTCA
MtST1     TTTTCCTTGA AGGTGGTGCT CAAATGCTCA TATGCCAGGT TGCAGTAGCA
VfSTP1    TTTTTCTCGA AGGTGGTGTT CAAATGCTTA TTTGTCAGGT TGCAGTTGCA
RcSTC     TTTTCCTTGA GGGTGGAGTT CAAATGTTGA TTTGCCAGGC AATTGTTGCA
VvHT      TTTTCCTTGA GGGTGGCACT CAAATGCTCA TATGTCAGGT TATTGTGGCA
AtSTP1    TCTTTCTTGA AGGTGGTACA CAAATGCTTA TATGCCAGGC TGTGGTTGCA
```

Figure 5.1. (continue)

```
          1201                                                    1250
LeHT3     ATAATCTTGG GACTCAAATT TGGAAGTGA. ........CA AGGAGCTATC
LeHT1     ATTTTGATAG CAATAAAGTT TGGAGTAAAT GGAACTCCAG GGGAATTACC
NtMST1    ATTTGCATAG CTATAAAATT TGGAGTGAAT GGAACTCCAG GGGATTTACC
MtST1     GCTGCAATTG GGGCCAAATT TGGAACAAGT GGAAACCCTG GTAATTTACC
VfSTP1    GTTTCAATTG CGGCCAAGTT TGGAACAAGT GGAGAACCTG GTGATTTACC
RcSTC     GCCTGCATTG GTGCTAAGTT TGGAGTAGAT GGAGCTCCCG GTGACTTGCC
VvHT      ACGTGCATTT TGGTTAAATT CGGAGTGGAT GGAGAACCTT GGTGCTTGCC
AtSTP1    GCTTGCATAG GGGCCAAGTT TGGGGTAGAC GGGACCCCTG GTGAGCTACC

          1251                                                    1300
LeHT3     AAGAGGTTAC TCGATTATAG TAGTTGTTTT CATTTGCCTC TTTGTAGCGG
LeHT1     AAAATGGTAT GCAATAGTGG TTGTGATATT CATTTGTGTA TATGTTGCTG
NtMST1    AAAGTGGTAC GCGATAGTAG TGGTGATATT CATCTGTGTT TATGTAGCTG
MtST1     AGAATGGTAT GCTATAGTAG TTGTGCTCTT CATTTGCATT TACGTAGCAG
VfSTP1    AAAGTGGTAT GCTATAGTAG TTGTGCTTTT CATATGCATT TACGTTGCTG
RcSTC     ACAATGGTAT GCAGTCGTTG TGGTGCTTTT CATTTGCATT TACGTATCTG
VvHT      CAAGTGGTAT GCCATAGTTG TGGTGCTGTT CATTTGCGTC TATGTTTCAG
AtSTP1    AAAGTGGTAT GCTATAGTGG TTGTAACGTT CATTTGCATC TATGTGGCGG

          1301                                                    1350
LeHT3     CGTTTGGATA CTCATGGGGG CCTCTTGGAT GGACCGTGCC AAGTGAAATT
LeHT1     GATTCGCTTG GTCATGGGGT CCTCTTGGAT GGCTCGTACC TAGTGAAATT
NtMST1    GATTTGCTTG GTCCTGGGGA CCTCTAGGAT GGTTGGTACC TAGTGAAATT
MtST1     GATTTGCTTG GTCATGGGGT CCTCTTGGTT GGTTGGTTCC TAGTGAGATT
VfSTP1    GATTTGCTTG GTCATGGGGT CCTCTTGGTT GGTTGGTGCC TAGTGAGATT
RcSTC     GATTCGCCTG GTCTTGGGGT CCCCTGGGAT GGCTGGTGCC AAGTGAAATC
VvHT      GGTTTGCATG GTCCTGGGGA CCTCTAGGTT GGTTGGTCCC TAGTGAAATT
AtSTP1    GTTTTGCGTG GTCGTGGGGC CCACTAGGGT GGTTAGTACC GAGTGAAATC

          1351                                                    1400
LeHT3     TTCCCTTTAG AGACGAGATC AGCAGGCCAA AGTATCACAG TTACTGTGAA
LeHT1     TTCCCACTGG AAATTCGATC AGCTGCACAA AGTATCAATG TCTCAGTGAA
NtMST1    TTCCCACTTG AAATTCGATC AGCTGCTCAA AGTATCAATG TTTCAGTGAA
MtST1     TTCCCATTGG AGATTCGTTC TGCAGCTCAA AGTGTAAACG TATCTGTGAA
VfSTP1    TTTCCATTGG AGATCCGTTC TGCTGCGCAG AGTGTCAACG TGTCGGTCAA
RcSTC     TTCCCACTGG AAATTCGGTC TGCTGCACAA AGTGTGAATG TTTCTGTCAA
VvHT      TTCCCCCTGG AAATCCGATC TGCTGCACAG AGTGTAAACG TCTCCGTTAA
AtSTP1    TTCCCGTTGG AGATAAGGTC GGCGGCGCAG AGTATCACCG TGTCCGTGAA

          1401                                                    1450
LeHT3     TTTGTTCTTC ACATTTGCGA TAGCACAGTC TTTCCTCTCA CTTTTATGTG
LeHT1     CATGATCTTC ACATTTGCAG TAGCACAAGT TTTCTTAACA ATGTTGTGTC
NtMST1    CATGATCTTC ACATTTATAG TGGCACAAGT ATTCTTGACA ATGTTGTGTC
MtST1     CATGCTTTTC ACCTTCTTAG TTGCACAAGT TTTCTTGATA ATGCTTTGTC
VfSTP1    CATGCTCTTC ACCTTCTTAG TTGCACAGAT TTTCTTGACC ATGCTTTGTC
RcSTC     CATGTTCTTT ACATTTGTAG TAGCTCAAGT ATTCCTGATA ATGCTTTGTC
VvHT      CATGTTTTTC ACATTCATCA TAGCCCAAAT CTTCTTAAAT ATGCTGTGTC
AtSTP1    CATGATCTTC ACGTTCATTA TCGCGCAAAT CTTCTTGACG ATGCTTTGTC
```

Figure 5.1. (continue)

```
          1451                                                    1500
LeHT3     CTATGAGGTT CGGGATTTTC CTGTTTTTCT CCTGTTGGAT TGCTGTCATG
LeHT1     ATTTGAAGTT TGGATTGTTT CTGTTTTTCG CCTTCTTTGT GGTGATTATG
NtMST1    ATTTGAAGTT TGGATTGTTC CTCTTCTTTG CATTCTTTGT TGTGATTATG
MtST1     ACATGAAGTT TGGTTTGTTC CTCTTCTTTG CCTTCTTCGT TTTGGTGATG
VfSTP1    ACATGAAGTT TGGATTGTTC CTCTTCTTTG CCTTCTTTGT GGTGGTGATG
RcSTC     ATTTGAAGTT TGGGCTATTC ATCTTCTTTT CATTCTTTGT GTTGATAATG
VvHT      ACATGAAGTT TGGTTTGTTC CTCTCCTTTG CCTTCTTTGT GGTGGTGATG
AtSTP1    ATTTGAAGTT TGGGTTATTC CTTGTTTTCG CCTTTTTCGT GGTGGTGATG

          1501                                                    1550
LeHT3     ACGATATTCA TCTATCTGTT CTTGCCTGAA ACGAAGGGAG TTCCGATTGA
LeHT1     ACTGTGTTCA TATACTTCTT CTTGCCTGAG ACGAAAAATA TTCCGATAGA
NtMST1    ACTGTCTTCA TTTACTTCTT CTTGCCTGAG ACAAAGAATA TCCCAATTGA
MtST1     TCAATCTATG TATTCTTCTT ATTGCCTGAA ACTAAAGGAA TACCAATTGA
VfSTP1    ACAATTTATA TATACACTAT GCTGCCTGAA ACTAAGGGAA TACCAATTGA
RcSTC     TCCATCTTCG TGTACTACTT CTTGCCTGAG ACAAAAGGCA TCCCAATTGA
VvHT      TCCTTCTTCA TTTACTTCTT CTTGCCTGAG ACCAAAGGCA TCCCAATTGA
AtSTP1    TCGATCTTTG TATACATTTT CTTGCCGGAG ACGAAAGGGA TTCCGATAGA

          1551                                                    1600
LeHT3     AGAGATGATG CGTCTTTGGG AAAAGCATTG GTTCTGGAAG AAGATCGTTT
LeHT1     AGAGATGGTG ATTGTGTGGA AAGAACATTG GTTCTGGTCT AAGTTCATGA
NtMST1    AGAGATGGTG ATTGTGTGGA AAGAGCATTG GTTCTGGTCT AAGTTCATGA
MtST1     AGAGATGGAC AGAGTTTGGA AATCACATCC CTTCTGGTCT AGATTTGTTG
VfSTP1    AGAGATGGAT AGAGTATGGA AATCACACCC ATATTGGTCT AGATTTGTTG
RcSTC     AGAGATGGGC CAAGTATGGA AGCAACACTG GTACTGGTCA AGATATGTTG
VvHT      AGAGATGGCT GAAGTATGGA AAAGTCACTG GTTCTGGTCC CGGTATGTCA
AtSTP1    GGAGATGGGT CAAGTGTGGA GGTCACACTG GTATTGGTCA AGGTTTGTGG

          1601                                                    1650
LeHT3     CGGAGGATCA ACAAGTTAAA AACACCAATG GACTCAACCA TGCTTGAGGT
LeHT1     CTGAAGTTGA TTATCCTG.. GAACTAGGAA TGGAACTGCT GTTGAAATGG
NtMST1    CTGAAGTGGA CTATCCTG.. GAACTAGGAA TGGAACAAGT GTTGAAATGT
MtST1     AACATGGTGA TCATGGCAAT GGTGTTGAGA TGGGAAAGGG AGCTCCTAAA
VfSTP1    AACATG.... ..ATGACAAT GGTGTTGAGA TGGCCAAGGG AGGTGTTAAA
RcSTC     TGGATGAAGA TTATCCTAAT GG...AGGGC TTGAAATGGG CAAGGAAGGC
VvHT      ACGATGGTTC TTACAGCGGC GT...CGAAC TGGTCAAGGA AAACTACCCT
AtSTP1    AGGATGGTGA GTATGGGAAT GCGCTTGAGA TGGGCAAGAA CAGTAACCAA

          1651                                                    1700
LeHT3     AGGTACATTG TTAACAAGGT GAAAGGTGAA TAGAGAAGGC TATCTTGTTC
LeHT1     CTAAAGGGGG TGCTGGTTAC AAAATTGTAT GACTTTAGTT TGGGTTTTTA
NtMST1    CAAAAGGGAG TGCTGGTTAC AAAATAGTAT GACCTAATTA AAG......A
MtST1     AATGTGTAAT TATTATTATT AGTCTTCATT TTATTTTATT TTCATTATTA
VfSTP1    AATGTATAAT TTATCAGTCT TGCTTTTATT TTTATTTTTG TTACTAATTA
RcSTC     CGAATTCCAA AGAATGTGTA ACTAAGCTTG GTGGATTTAT TTTAGCTTAT
VvHT      GTTAAGAATG TATGA~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    GCTGGAACGA AGCATGTTTG ATTTATCATT GTTTTTAATG AGAGTTTTAA
```

Figure 5.1. (continue)

```
          1701                                                      1750
LeHT3     ATGTAGAATT TTGGTCTTCT ATATGTATAT GTAAAAGTGT GAACTTTTTA
LeHT1     AATTTTTATT TGTTGTTTGT ATAATGTTGT AGTGGGGATG ATATTGATAA
NtMST1    AGATGTTTGG ATTTATTTTA ATTTTATTTG TTGTTGTATA ATGTTTTAGT
MtST1     ATTAGTTTTA TTGGTGAAAC ACTAACTATT GGTGTCAACC TCAAGTATCA
VfSTP1    GTTTTATTTT TATTTTGCTT GTGAAGCTTG GTATTGAATT GTCCTTAAAA
RcSTC     ATTTTGAACC CTTTTCCCTT TTCTTCTTAC CTTTTGGGAT TAGCAGCAGT
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    GAAAGAAAGA AAAAAGATTT GTAATTTCTA ATGTCGTAAA GGAAAAAGTG

          1751                                                      1800
LeHT3     CTTTTAAGAA TGTCACTTTC TATGAATACA CGAGTGAAAT GATATTTTGA
LeHT1     TTATTAATTA GATTTGATTG AAACTGTTTC TATTGTTTAC TTTTGCATAG
NtMST1    GGGATGATAT TGTTAGATTT GAACTGTTTC TCTTGTCTCT ATTGCATAGA
MtST1     AATGTAATGA AATTGCACTT CAAAATTACG GGATTATTTT TCTCAAAAAA
VfSTP1    AAAAAAAAAA AAA~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
RcSTC     ACATACTAAT GATTTCAATA TCAAAAATTA TGTGGAAATT TTT~~~~~~~
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    TATTAGCCTA GATATTTATT GGTGTTTATA TAATTCAATA CCACATGAAG

          1801                                                      1850
LeHT3     TAAAAAAAAA AAAAAA~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     AAAAAAATAC ATAACTTTGT TCAATAGAAA ATTTGGCAAA GCAACTGTGT
NtMST1    AAATACATAC TTTACTGTGT TCAATTTCAA CCTCTCAACA ATTAATATAA
MtST1     AAAAAAAAAA AAAAAAAAAA AAAAAA~~~ ~~~~~~~~~~ ~~~~~~~~~~
VfSTP1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
RcSTC     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    AAATTATGCA TATGATTCTT CGTTAATTGT CCGTAATTGT TATACTCTTT

          1851                                                      1900
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     AGCCTTTTGT TTGTTTTGTT GGATGTACTA TTTAGTAGTT CTAGTCTTTT
NtMST1    GATTTGTCAG AGCAAAA~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
MtST1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
VfSTP1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
RcSTC     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    ACTTAAACCA AGTGTTTTCT CTTTGAAAAA AAAAAAAAAA AAAAAA~~~~

          1901                                                      1950
LeHT3     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
LeHT1     AGTGTAAGAT CTTTATTATA AAAAATTATG TTCATAAGCT GTATAAAAAA
NtMST1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
MtST1     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
VfSTP1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
RcSTC     ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
VvHT      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
AtSTP1    ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.1. (continue)

```
          1951          1964
LeHT3     ~~~~~~~~~~ ~~~~
LeHT1     AAAAAAAAAA AAAA
NtMST1    ~~~~~~~~~~ ~~~~
MtST1     ~~~~~~~~~~ ~~~~
VfSTP1    ~~~~~~~~~~ ~~~~
RcSTC     ~~~~~~~~~~ ~~~~
VvHT      ~~~~~~~~~~ ~~~~
AtSTP1    ~~~~~~~~~~ ~~~~
```

Figure 5.1. (continue)

5.2.7 Full-length cDNA isolation

Five μg of total RNA of 'Sunset' mature green fruit was dephosphorylated with calf intestinal phosphatase (CIP) and then decapped with tobacco acid pyrophosphatase (TAP) following the manufacturer procedure (GeneRacer™ Kit - Invitrogen, Carlsbad, California). The full-length decapped mRNA was ligated with GeneRacer™ RNA Oligo that contained the specific sequence for the GeneRacer™ 5' primer and 5'-nested primer. The ligated full-length mRNA was reverse-transcribed with SuperScript™ III RT and the GeneRacer™ Oligo dT primer containing specific sequence for GeneRacer™ 3' and 3'-nested primers to make first strand cDNA.

To amplify the cDNA 5' end, five gene specific primers (GSP) for papaya hexose transporter were designed from the 5' end of the known sequence of the first cDNA fragment as required by the GeneRacer™ Kit (Table 5.1). The primers, GSP5' and GSP5'-nested, were amplified in the following thermocycling regime: 95°C for 7 min, followed by 95°C for 1 min, 58°C for 1 min, 72°C for 30 sec (for 31 cycles), followed by the final extension period of 72°C for 7 min using the PTC1148 in the MJ Mini™ Gradient Thermal Cycler (Bio-Rad laboratories, Inc., Hercules, CA).

Papaya invertase forward (PpyInvF234) and reverse primers (PpyInvR787) were used as positive controls. Sterile water was used instead of the cDNA template in the negative control. The amplification system was as follow:

Component	Volume (µl)
Sterile, deionized water	5.5
ImmoMix (BioLine USA Inc., Randolph, MA)	12.5
cDNA template	2.0
10 µM GeneRacer 5' or 5'-nested primers (or GeneRacer 3' or 3'-nested primers)	3.0
10 µM GSP	2.0
Total volume	25.0

To amplify the cDNA 3' end, six potential gene specific primers (GSP) for papaya hexose transporter were designed from the 3' end of the known sequence of the first cDNA fragment as required by the GeneRacer™ Kit (Table 5.2). The primers were used as GSP3', and GSP3'-nested. The thermocycling regime was as follows: 95°C for 7 min, followed by 95°C for 1 min, 60°C for 1 min, 72°C for 30 sec (for 35 cycles), followed by the final extension period of 72°C for 7 min using the PTC1148 in the MJ Mini™ Gradient Thermal Cycler (Bio-Rad laboratories, Inc., Hercules, CA). The PCR products were purified from 1% agarose (1xTAE) gel using QIAEX II Gel Extraction Kit (QIAGEN).

5.2.8 RT-PCR, cloning and sequence analysis of the fuli-length cDNA

First strand cDNA was synthesized from 10 µg of total RNA of the mature green 'Sunset' papaya fruit tissue (118 DAA) using Oligo (dT) primer following the instruction of StrataScript™ Reverse Transcriptase (Stratagene, La Jolla, CA, Cat. No.600085). Two specific primers were designed from conserved regions of papaya hexose transporter cDNA sequences obtained from steps 5.2.9 and 5.2.10. The sequences of primers were as follow: forward primer (PpyHxTF1080) 5'- GGA TAC GAT ATT GGG ATC TCA GG- 3' (Tm 54.3°C) and reverse primer (PpyHxTR2743) 5' -CGA GAG TTT TTC GTA ATT CC- 3' (Tm 48.9°C). Papaya gene specific primers for acid invertase were used as positive control. The sequences of the invertase primers

were as follow: forward primer (PpyInvF234) 5'- TGG ATT AAC GAC CCA AAT GC - 3' (Tm 53°C) and reverse primer (PpyInvR787) 5'- AAT CTG GGC ATT CCC ACA TA - 3' (Tm 53°C). Sterile water was used instead of the cDNA template in the negative control. The amplification system was as follow:

Component	Volume (µl)
Sterile, deionized water	8.5
ImmoMix (BioLine USA Inc., Randolph, MA)	12.5
cDNA template	2.0
25 µM Forward primer	1.0
25 µM Reverse primer	1.0
Total volume	25.0

The thermocycling regime used was: step 1 initial denaturation at 95°C for 7 min; step 2 denatured at 95°C for 1 min; step 3 annealed at 47/50.2/53.4°C for 1 min; step 4 extension at 72°C for 1 min, PCR at steps 2 to 4 was repeated for 35 cycles, followed by step 5 final extension at 72°C for 7 min using the Eppendorf Mastercycler® personal 5332. The PCR products were 553 and 1,663 base pairs (bp) in size, for invertase and hexose transporter, respectively, were purified from a 1% agarose (0.5xTBE) gel using QIAEX II Gel Extraction Kit (QIAGEN).

Table 5.1. Sequences of primers used in nested PCR to amplify the 5' end of the gene as described in methods. Primers are in the 5' to 3' direction.

Primer name	Primer sequence (5'→3')	Tm (°C) estimated/true
GeneRacer 5' primer	CGA CTG GAG CAC GAG GAC ACT GA	74
GeneRacer 5' nested	GGA CAC TGA CAT GGA CTG AAG GAG TA	78
PpyHxTR1103	CCC TGA GAT CCC AAT ATC GTA TCC	72/56.3
PpyHxTRn1309	ACG TAA GTG TCG GAC TGT CGT ACT GA	78/60.7
PpyHxTR1377	CTT TCT TGT CAC CGT CGC CGC AAC CA	82/65.2
PpyHxTR1403	AAC AGC ATC GAC AGT TTC CGA CCG AAC T	84/63.4
PpyHxTR1535	GAG AGG TAG AGT GGT ACA GAC TGA TTG G	84/59.3

Table 5.2. Sequences of primers used in nested PCR to amplify the 3' end of the gene as described in methods. Primers are in the 5' to 3' direction.

Primer name	Primer sequence (5'→3')	Tm (°C) calculated/true
GeneRacer 3' primer	GCT GTC AAC GAT ACG CTA CGT AAC G	76
GeneRacer 3' nested	CGC TAC GTA ACG GCA TGA CAG TG	72
PpyHxTF1352	TGG TTG CGG CGA CGG TGA CAA GAA AG	82/65.2
PpyHxTF1645	AAA GGA GGT TGG GGA TGG AGA CTG A	76/61.5
PpyHxTF1657	TGG GGA TGG AGA CTG AGC TTA GGA	74/61.0
PpyHxTF1721	TAA TCC TCC CCG ATA CAC CCA ACT C	76/59.7
PpyHxTF2259	GTG GCC GGA TTT GCG TGG TCT TG	74/63.2
PpyHxTF2301	GTG CCG AGT GAA ATC TTC CCT CTT G	76/59.6

The purified DNA fragment was ligated into pGEM®-T Easy vector (Promega, Madison, Wisconsin) at the ratio of 3:1 for insert DNA: vector. The ligated DNA/vector (5 µl) was transformed into 50 µl One Shot® TOP 10 competent cells (Invitrogen, Carlsbad, California) as stated in step 5.2.8. The plasmid DNA was purified using QIAprep Spin Miniprep Kit (QIAGEN, Valencia, California). The amount of the plasmid DNA was quantified by reading the absorbance at 260 nm. The quality of the plasmid DNA was determined by cutting with Not1 restriction enzyme (NEBcutter, New England BioLabs) overnight at 37°C and then it was run on 1% agarose/ 0.5xTBE electrophoresis gel.

Plasmid DNA was sequenced by a 96-capillary 3730xl DNA Analyzer (Applied Biosystems, Hitachi, Japan) at the Center for Genomics, Proteomics, and Bioinformatics Research Initiative, University of Hawaii at Manoa. The full papaya hexose transporter sequence was searched in BLASTn against NR database of NCBI to confirm that the PCR products have homology to known monosaccharide transporters.

5.2.9 Non-radioactive RNA probe preparation

DIG-labeled RNA probes were made using the *in-vitro* transcription labeling according to the manufacturer procedure for 'DIG RNA Labeling Kit (SP6/T7)' (Roche Applied Science, Indianapolis, Illinois). The plasmid DNA was extracted from the One Shot® TOP 10 competent cells (Invitrogen, Carlsbad, California) (from step 5.2.8) according to the HiSpeed Midi kit's instructions (Qiagen, Valencia, California). The pGEM®-T Easy vector containing the estimated 1,444 bp ORF of the hexose transporter protein and SP6 RNA promoter upstream of ORF was used for RNA probe preparation. The plasmid was linearized with the restriction enzyme Nco1. The *in-vitro* transcription labeling was performed from one µg of linearized plasmid DNA template and DIG-11-UTP labeling nucleotide mixture using SP6 RNA polymerase. The mixture was incubated at 37°C for 2 h and 0.2M EDTA was added to stop the reaction. The yield of newly transcribed DIG-labeled-RNA probe of approximately size of 700 nucleotides was determined by serial dilution according to the manufacturing protocol 'DIG RNA Labeling Kit

(SP6/T7)' (Roche Applied Science, Indianapolis, Illinois).

5.2.10 Northern blot analysis

Northern blot analysis was performed according to the method of Fourney *et al.* (1988). 30 μg of total RNA from each stage of 'Sunset' papaya (from step 5.2.11) was run on 1% agarose gel containing 0.66M formaldehyde. The RNA gel was rinsed with transferred buffer, 10X SSC (1.5M NaCl, 150 mM Sodium citrate, pH7.0) twice for five min each prior to blotting. The RNA was transferred to the positive charge nylon membrane using 10X SSC with VacuGene XL Vacuum blotting system (GE Healthcare Bio-Sciences Corp, Piscataway, New Jersey). The RNA was fixed onto the membrane by UV-cross linking and the membrane was briefly rinsed with distilled water and allowed to air dry. The membrane was prehybridized with hybridization buffer, DIG Easy Hyb, in a hybridization tube with gentle rotation in the hybridization incubator at 68°C (for RNA-RNA binding) for 1 h following the guideline of DIG Northern Starter Kit (Roche Applied Science, Indianapolis, Illinois). The DIG-labeled RNA probe was denatured by boiling for 5 min and rapidly cooled in ice/water. 100 ng mL^{-1} of denatured DIG-labeled RNA probe was added into the pre-warm hybridization buffer, mixed, and then incubated at 68°C overnight with gentle agitation. The membrane was stringency washed at room temperature with excess amount of 2X SSC, 0.1% SDS twice for 5 min under constant vigorous agitation in a RNase free tray, and then washed with 0.1XSSC, 0.1% SDS twice for 15 min at 68°C. Detection was performed as described in the manufacturer's instruction of the DIG High Prime DNA Labeling and Detection Starter Kit II (Roche Applied Science, Indianapolis, Illinois) using CSPD chemiluminescent as substrate for alkaline phosphatase. The membrane was exposed on X-ray film and incubated at 37°C for 30 min following by room temperature for two hours for the first time. After 30 min incubation at 37°C, multiple exposures were taken to achieve the desired signal strength.

5.3 Results

5.3.1 Papaya fruit development during maturity and ripening

'Sunset' papaya fruit fresh weight remained constant during the period 104 to 153 DAA (Figure 5.2). Seed and flesh color started to develop 104 DAA, while skin color started to change at 118 DAA. Seed color reached 100% black at 132 DAA, one week prior to the flesh becoming fully yellow (139 DAA), and the skin color was 28 to 35% yellow (Figure 5.3). The pattern of TSS accumulation was similar to DW accumulation. TSS content continuously increased from 104 DAA, started one week prior to DW accumulation (111 DAA) and reached a maximum at 153 DAA (Figure 5.4).

5.3.2 First partial *CpHT1* cDNA amplification and sequencing analysis

The first cDNA of papaya hexose transporter gene (*CpHT1*) was isolated by PCR from 'Sunset' papaya fruit using the degenerate primers HxTF486 and HxTR1076 at different annealing temperatures between 45 and 56°C (Figure 5.5). The PCR product of 590 bp in length was cloned into pGEMT®-easy vector for amplification in One Shot® TOP10 competent cells and sequencing. BlastN search indicated that the sequence was matched the sequence of hexose transporters in many plants including grape berry *VvHT* (Y09590 and AJ001061), poplar *tMST2.2* (AJ698938), castor bean *RcSTC* (L08196), barrel medic *MtST1* (U38651), *Arabidopsis AtSTP1* (X55350), olive *OeMST2* (DQ087177), fava bean (Z93775), soybean *GmSTP1* (AJ563365), and tobacco *NtMST1* (X66856).

5.3.3 Amplification and sequence analysis of the 5' ends of the first partial *CpHT1* cDNA

Five gene specific reverse primers (Table 5.1) designed from the first partial *CpHT1* sequence were used to amplify the 5' end with the 5' and 5' nested primers. Only one PCR product was amplified from PpyHxTR1377 and GeneRacer 5' primer (Figure 5.6) and the sequence did not match known hexose transporters.

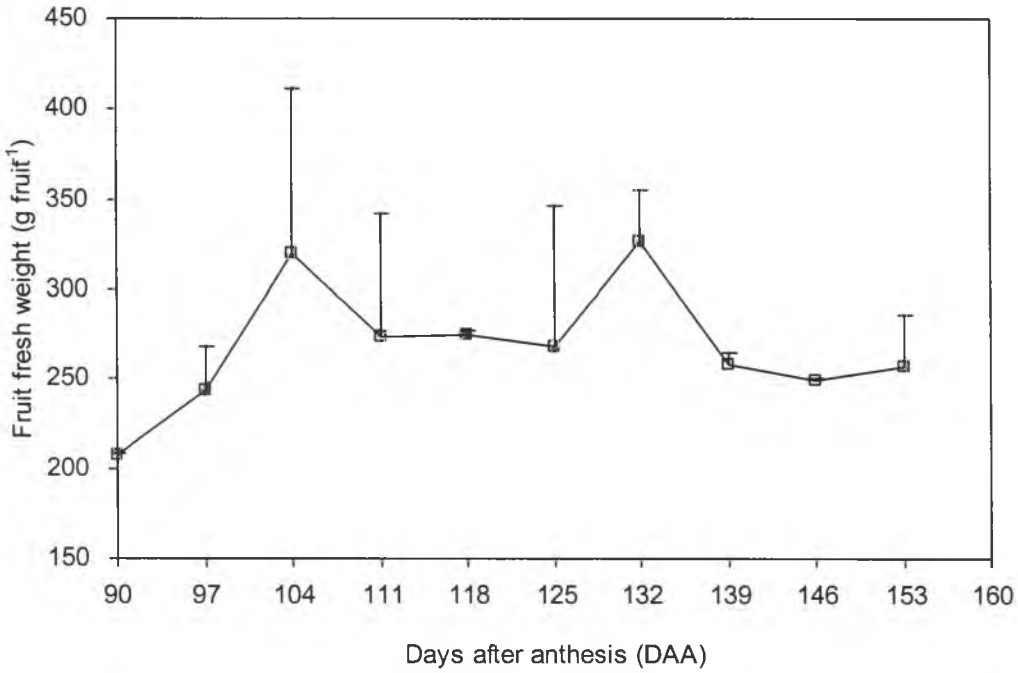

Figure 5.2. 'Sunset' papaya fruit growth from fruit harvested between 90 and 153 DAA in 2006. Mean ±SD, n = 3.

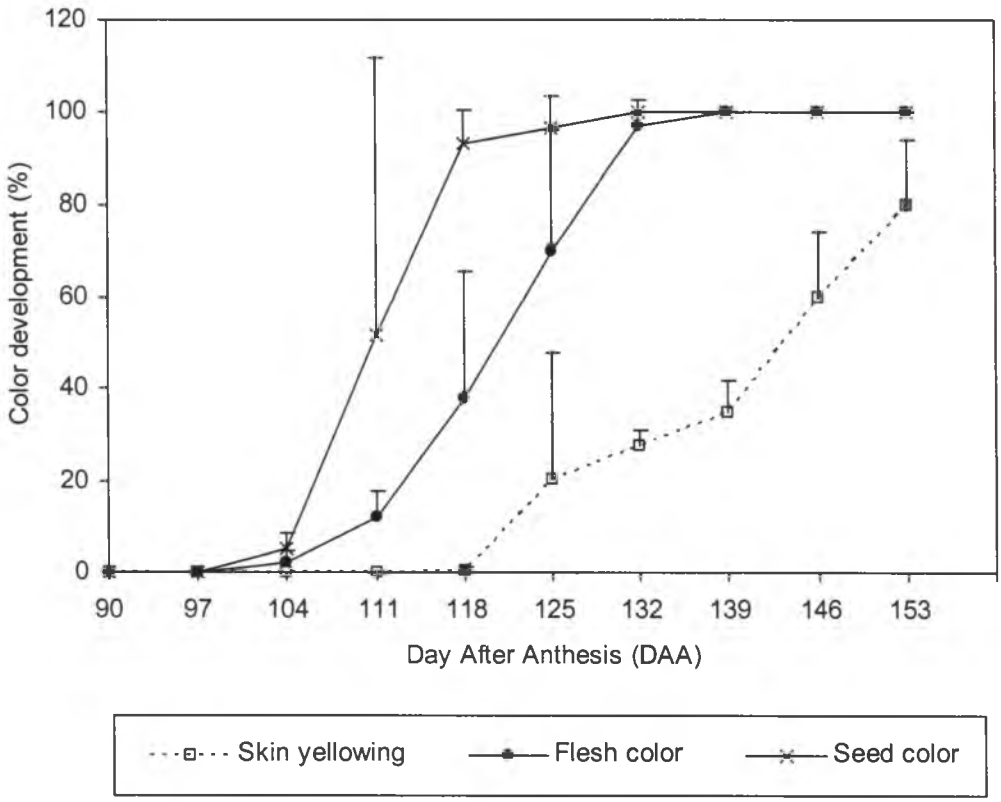

Figure 5.3. 'Sunset' papaya skin, flesh and seed color development of fruit harvested in 2006. Mean ±SD, n = 3.

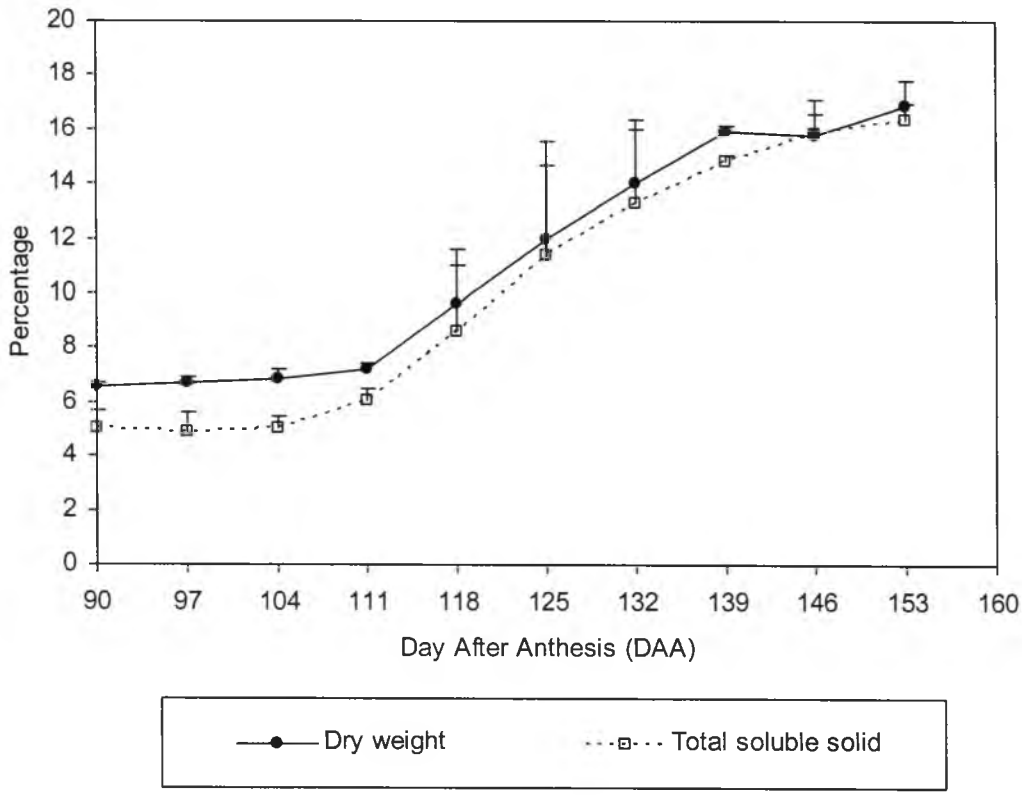

Figure 5.4. Dry weight and totai soluble solids of 'Sunset' papaya flesh from fruit harvested between 90 and 153 DAA in 2006. Mean ±SD, n = 3.

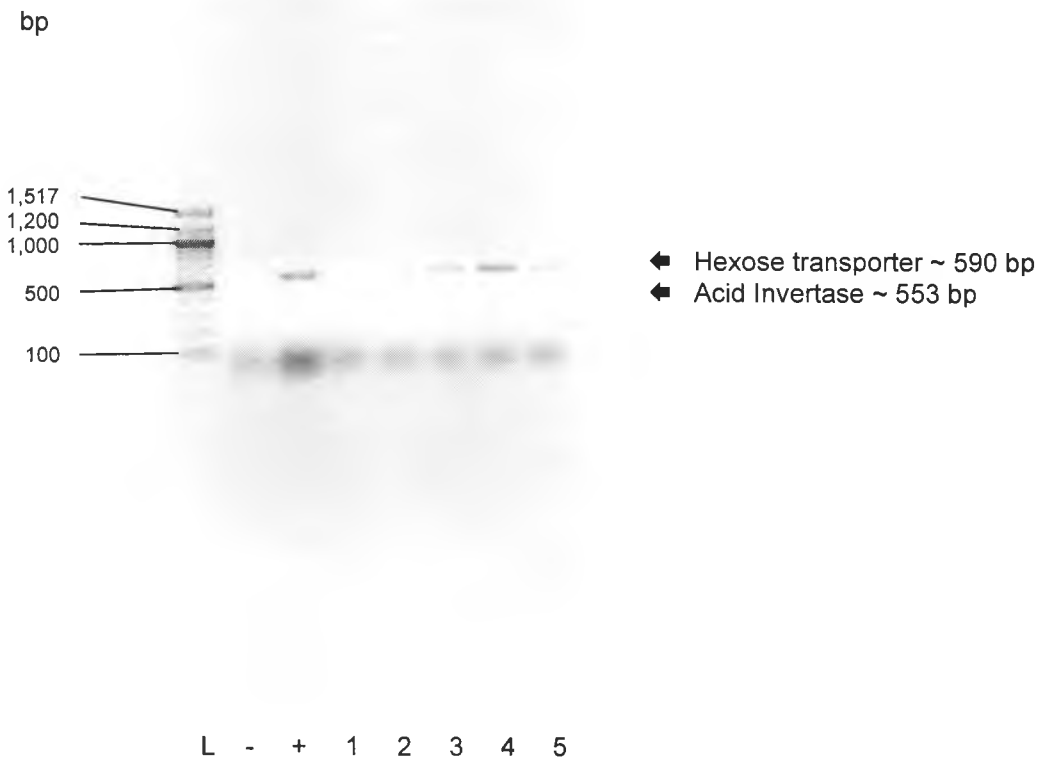

Figure 5.5. The first 590 bp 'Sunset' papaya hexose transporter cDNA (*CpHT1*) amplified from a degenerate forward primer (HxTF486) and reverse primer (HxTR1076) at different annealing temperatures. The electrophoresis was run on 1% agarose gel with 1xTAE and 0.5 µg/ml ethidium bromide buffer.

L = 500 ng of 100 bp DNA ladder

- = Negative control using sterile water instead of cDNA template

+ = Positive control using invertase forward and reverse primers

1 = Hexose transporter PCR product annealing at 45.4°C

2 = Hexose transporter PCR product annealing at 48.0°C

3 = Hexose transporter PCR product annealing at 50.7°C

4 = Hexose transporter PCR product annealing at 53.5°C

5 = Hexose transporter PCR product annealing at 56.0°C

5.3.4 Amplification and sequence analysis of the 3' end of the first partial *CpHT1* cDNA

Six gene specific forward primers (Table 5.2) were designed from the first partial *CpHT1* sequence and used to amplify the 3' end. Gel electrophoresis indicated no single band from any primer pairs. At least five PCR products were amplified from PpyHxTF1645 and GeneRacer 3' primer (Figure 5.7). However, only the sequence of the 1.2 kb long amplicon matched hexose transporters from other plants and this fragment had 100% overlap with the first partial *CpHT1* cDNA. To confirm that the first 590 bp fragment and the second 1.2 kb *CpHT1* cDNAs were of the same gene, a gene specific forward primer, PpyHxTF1080, and two gene specific reverse primers, PpyHxTR2322 and PpyHxTR2276, were designed from the 590 bp *CpHT1* (Figure 5.5) and the second 1.2 kb *CpHT1* (Figure 5.7), respectively, to amplify the connected amplicon between the first two PCR products. Two 1.2 kb long PCR products from two pairs of primers were purified and cloned into pGEMT®-easy vector for amplification in One Shot® TOP10 competent cells (Figure 5.8). The comparison between these two sequences and the first two sequences indicated that they were the same *CpHT1* cDNA and the sequence matchedthat of hexose transporters from other plants. However, the sequence of *CpHT1* did not contain the poly A tail.

A fourth amplification was run using ImmoMix (BioLine USA Inc., Randolph, MA) with various gene specific forward primers (Table 5.2) and GeneRacer 3' and 3'-nested primer to amplify the poly A tail of *CpHT1*. No PCR product was amplified by PpyHxTF1645/GeneRacer 3' primer (Figure 5.9) whereas at least five products were amplified by TaqPfx high-fidelity polymerase when used the same primer pairs (PpyHxTF1645/GeneRacer 3' primer) (Figure 5.7). Sequence analysis indicated that the PCR products amplified from PpyHxTF1721/ GeneRacer 3' primer (clone 11, band A), PpyHxTF2259/GeneRacer 3' primers (clone 9), PpyHxTF2301/GeneRacer 3'-nested primers (clone 10) and PpyHxTF1721/GeneRacer 3'-nested primers (clone 11, band A) were the same *CpHT1* cDNA and matched the genomic *CpHT1* DNA. However, only clone 9s contained a poly A tail (Figure 5.9). Sequence analysis indicated that clone 9-7 was the only clone one out of seven clones (9-1, 9-5, 9-6, 9-7, 9-8, 9-10 and 9-11) from the PCR product amplified by GeneRacer 3' primer and PpyHxTF2259 that contained the actual

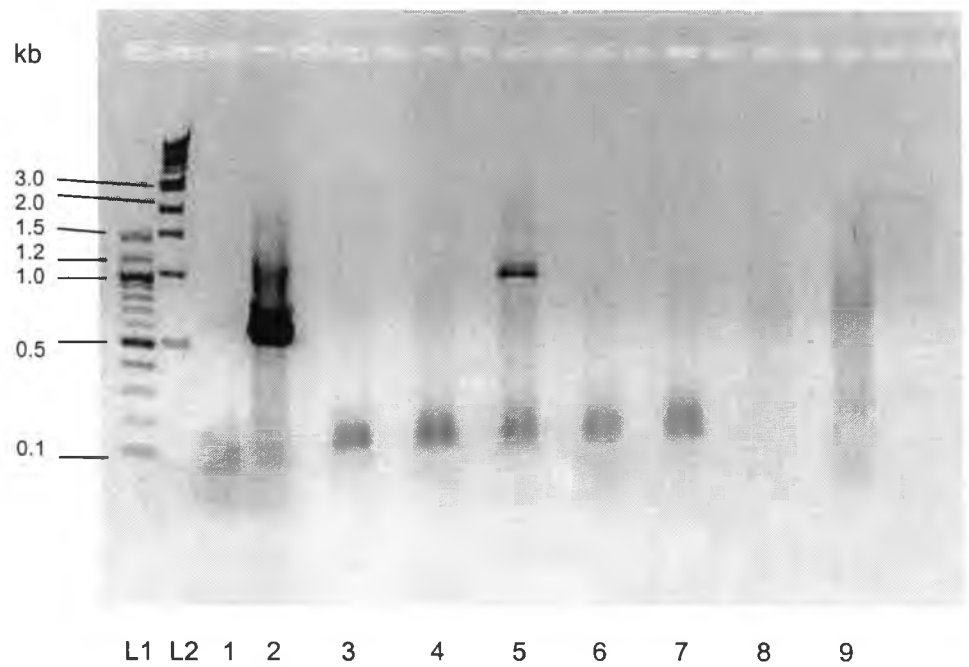

Figure 5.6. The 5' extended papaya hexose transporter amplicon from various gene specific reverse primers designed from the beginning of the first sequence and GeneRacer 5' and 5'-nested primers. The electrophoresis was run on 1% agarose gel with 0.5xTBE buffer.

L1 = 500 ng of 100 bp DNA ladder

L2 = 500 ng of 1.0 kb DNA ladder

1 = Negative control using sterile water instead of cDNA template

2 = Positive control using invertase forward and reverse primers

3 = PpyHxTR1535 and GeneRacer 5' primer

4 = PpyHxTR1403 and GeneRacer 5' primer

5 = PpyHxTR1377 and GeneRacer 5' primer

6 = PpyHxTRn1309 and GeneRacer 5' primer

7 = PpyHxTR1103 and GeneRacer 5' primer

8 = PpyHxTR1103 and GeneRacer 5' nested primer

9 = PpyHxTRn1309 and GeneRacer 5' nested primer

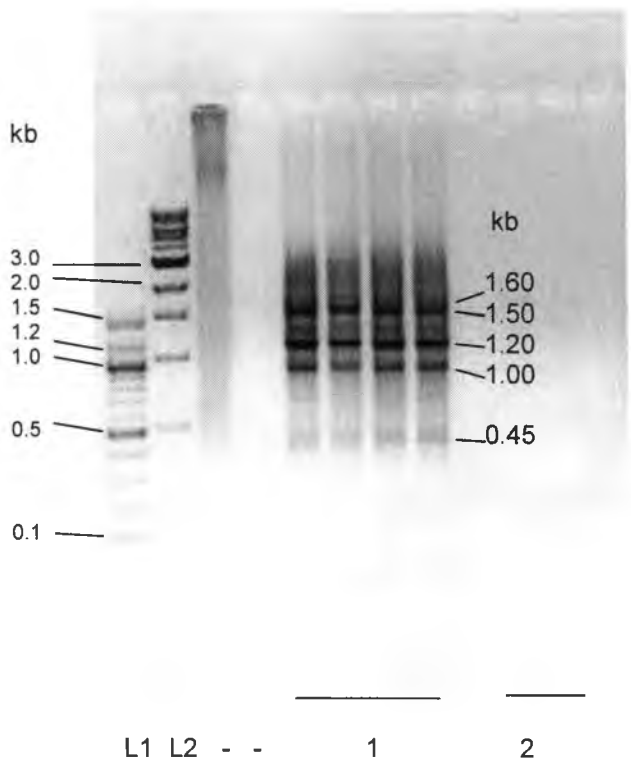

Figure 5.7. The 3' extended papaya hexose transporter cDNA amplified from the gene specific forward primers, designed from the sequence of the first clone, and GeneRacer 3' and 3'-nested primers using TaqPfx polymerase. The electrophoresis was run on 1% agarose gel with 0.5xTBE buffer.

L1 = 500 ng of 100 bp DNA ladder

L2 = 500 ng of 1.0 kb DNA ladder

- = Negative control using sterile water instead of cDNA template

1 = PpyHxTF1645 and GeneRacer 3' primer

2 = PpyHxTF1721 and GeneRacer 3' nested primer

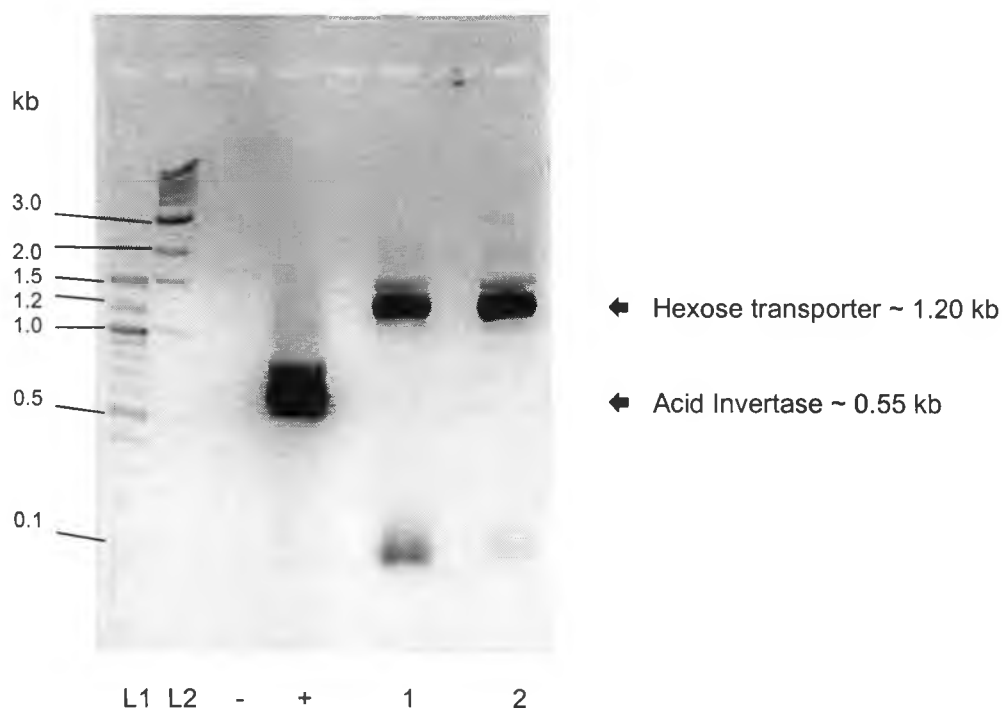

Figure 5.8. The 1.2 kb papaya hexose transporter cDNA amplicon from the gene specific forward primer (PpyHxTF1080) designed from the beginning of the first amplicon and two gene specific reverse primers (PpyHxTR2322 and PpyHxTR2276) from the end of the second amplicon to connect the sequence between the first two cDNAs. The electrophoresis was run on 1% agarose gel with 0.5xTBE buffer.

L1 = 400 ng of 100 bp DNA ladder

L2 = 300 ng of 1.0 kb DNA ladder

- = Negative control using sterile water instead of cDNA template

+ = Positive control using invertase forward and reverse primers

1 = PpyHxTF1080 and PpyHxTR2322 (expected size ~ 1,242 bp)

2 = PpyHxTF1080 and PpyHxTR2276 (expected size ~ 1,196 bp)

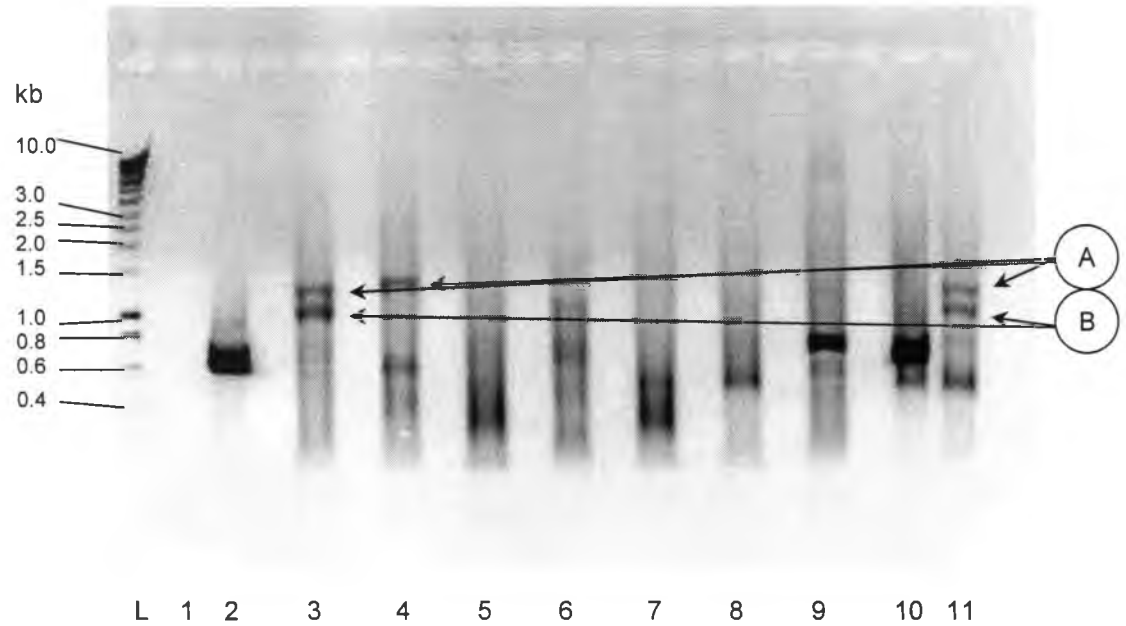

Figure 5.9. The second 3' extended papaya hexose transporter amplicon (a forth amplicon) from various gene specific forward primers designed from the end of the second sequence and GeneRacer 3' and 3' nested primers using ImmoMix. The electrophoresis was run on 1% agarose gel with 0.5xTBE buffer.

L = 720 ng Hyperladder I (Bioline)

1 = Negative control using sterile water instead of cDNA template

2 = Positive control using invertase forward and reverse primers

3 = PpyHxTF1721 and GeneRacer 3' primer

4 = PpyHxTF1657 and GeneRacer 3' primer

5 = PpyHxTF1645 and GeneRacer 3' primer

6 = PpyHxTF1352 and GeneRacer 3' primer

7 = PpyHxTF1645 and GeneRacer 3' nested primer

8 = PpyHxTF1352 and GeneRacer 3' nested primer

9 = PpyHxTF2259 and GeneRacer 3' primer

10 = PpyHxTF2301 and GeneRacer 3' nested primer

11 = PpyHxTF1721 and GeneRacer 3' nested primer

poly (A) additional site (T) located 154 nt downstream from the stop codon whereas the rests were 19 nt shorter.

5.3.5 Amplification and sequence analysis of the full length *CpHT1* cDNA

The final 1,663 bp cDNA was amplified using gene specific forward primer (PpyHxTF1080) designed from the beginning of the first partial 590 bp-amplicon, and the gene specific reverse primer (PpyHxTR2743) designed from the end of the last 558 bp-amplicon that contained the poly A tail (Figure 5.10) to confirm that all of the overlapped five amplicons (590, 658, 1,242, 840, and 558bp long, Figure 5.12) were the same gene (*CpHT1*). The 1,663 bp PCR product was cloned into pGEMT®-easy vector and amplified. To assure that the cDNA was inserted in the vector, the purified insert DNA/pGEMT®-easy vectors were excised with Not1 restriction enzyme (Figure 5.11) and the PCR products were shown to be the same *CpHT1* cDNA (Figure 5.12). Sequence analysis and the multiple comparisons between all partial and final papaya hexose transporter sequences and grape *VvHT* (Y09590) indicated that they were homologous (Figure 5.13 and 5.14). The combined sequence of the 1,663 bp (Figure 5.10) and the poly A tail of clone 9 (GR3F2259) and 11A (GR3F1721) (Figure 5.9) was screened with BLASTn. The comparison showed a high homology to hexose transporter sequence from a large group of monocots and dicots species from various tissues. Papaya hexose transporter gene showed the highest score (bits) to poplar (80% Identity, AJ698938), grape berry *VvHT* (82% Identity, Y09590), English walnut (*Juglans regia HT1*, 79% identity, DQ026508), castor bean *RcSTC* (L08196), tobacco *NtMST1* (X66856), barrel medic *MtST1* (U38651), grape *VvHT1* (AJ001061), and *Arabidopsis AtSTP1* (X55350), respectively.

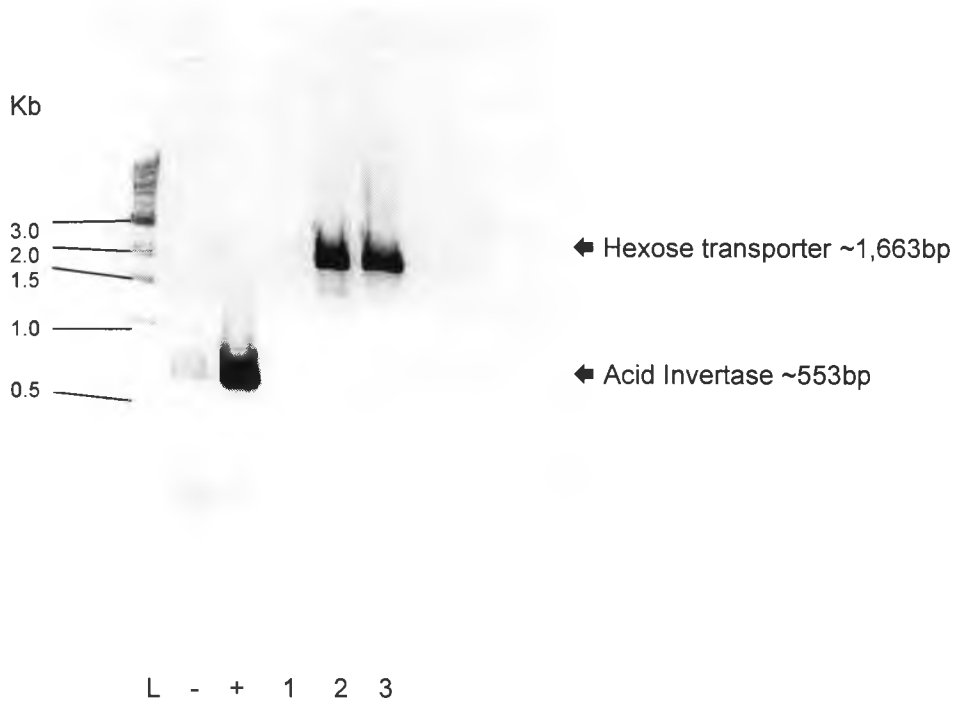

Figure 5.10 The final (fifth) 'Sunset' papaya hexose transporter cDNA amplicon from the gene specific forward primer (PpyHxTF1080) designed from the beginning of the first amplicon and the gene specific reverse primer (PpyHxTR2743) designed from the end of the forth amplicon at different annealing temperature indicated the same gene after all. The electrophoresis was run on 1% agarose gel with 1xTAE buffer.

L = 500 ng of 1kb DNA ladder

- = Negative control using sterile water instead of cDNA template

+ = Positive control using papaya invertase forward and reverse primers

1 = Hexose transporter PCR product annealing at 47.0°C

2 = Hexose transporter PCR product annealing at 50.2°C

3 = Hexose transporter PCR product annealing at 53.4°C

5.3.6 Structure of the full length *CpHT1* cDNA

The nucleotide sequence of the papaya hexose transporter cDNA (*CpHT1*) was identical to the sequence of the predicted exons of the papaya genome supercontig_1226. This supercontig also contained a motif similar to the hexose transporter gene of poplar *tMST2.2* (AJ698938) (Figure 5.15). The gene was composed of four exons and three introns. The translation initiation codon (ATG) was located at base 284. This ATG was more likely to represent the functional translation start codon, since it gave the complete open reading frame of 523 amino acids. The sequence between the ATG site and the attachment site of the poly A tail or the poly A site (T) was 2,979 bp long.

The 5' untranslated end contained both CAAT and TATA motifs. The single putative CCAAT motif was located – 211 nt upstream from start codon. Three putative TATA motifs were located – 199, – 119 and – 107 nt upstream from the start codon. All three TATA motifs covered 8 – 10 nt, however, they did not corresponding to the consensus sequence (C/G)TATA(T/A)A1–3(C/T)A. The initiator site (Inr), pyrimidine rich region, was located between -36 and -170 covering TATA-119 and TATA-107 boxes. A purine rich region was located between -1 and -35 upstream from the start point. The splicing consensus sequences AG/GT were conserved for the introns of *CpHT1*. 'CAG' codon was observed on both 5' and 3' splice sites of all three introns (Figure 5.15 and 5.16).

The 3' untranslated end of *CpHT1* contained a short sequence of poly T(U) tract, 1 nt downstream from the stop codon (TGA). The stop codon (TGA) was located at base 3,108. The sequence between start (ATG) and stop (TGA) codons was 2,825 bp long. The possible transcription termination signal, AAT(U)AAA in *CpHT1* was found 131 nt downstream from the termination codon (TGA) or -24 nt upstream from the poly A site (T), however, this signal was incomplete (AATAA), other AATAAAs were found after the poly A site (T) (Figure 5.15).

The phylogenic diagram and multiple comparison of hexose transporter genes between HTs of other plants and the full length *CpHT1* cDNA showed that *CpHT1* belongs to the same class with poplar *tMST2.2* (AJ698938), grape *VvHT* (Y09590), castor bean *RcSCP* (L08196), and peach fruit *PpSTP1* (AF367454) (Figure 5.17).

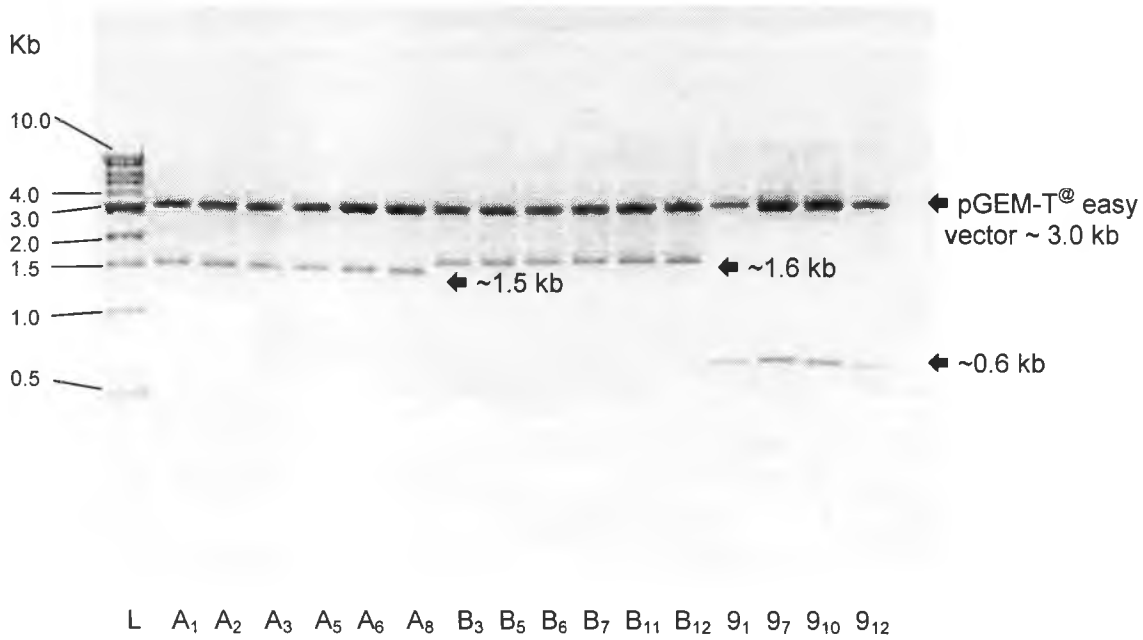

Figure 5.11. Not1 cut from papaya hexose transporter clones confirmed the insertion of the cDNA fragments within the pGEM-T® easy vectors (3 kb). The electrophoresis was run on 1% agarose gel with 0.5xTBE buffer.

L = 300 ng of 1.0 kb DNA ladder

A_1, A_2, A_3, A_5, A_6, A_8 = Clone A (PpyHxTF1080 and PpyHxTR2622)

B_3, B_5, B_6, B_7, B_{11}, B_{12} = Clone B (PpyHxTF1080 and PpyHxTR2743)

9_1, 9_7, 9_{10}, 9_{12} = Clone 9 (PpyHxTF2259 and GeneRacer 3' primer)

[HxTF486]-----590 bp---→[HxTR1076] (Figure 5.5)

[PpyHxTF1645]----658 bp----→[GeneRacer 3' primer] (Figure 5.7)

[PpyHxTF1080]----------1,242 bp----------→[PpyHxTR2322] (Figure 5.8)

PpyHxTF[1721]----- 840 bp---→[GeneRacer 3' primer or 3' nested primer]

(band A, Figure 5.9)

[PpyHxTF2259]—~600 bp→[GeneRacer 3' primer]

(Figure 5.9)

[PpyHxTF1080]----------1,663 bp----------→ [PpyHxTR2743]

(Figure 5.10)

Figure 5.12. Summary diagram showed the overlapped sequences of all amplicons and the final 1,663 bp long *CpHT1* cDNA.

Figure 5.13. The multiple comparisons between all partial *CpHT1* cDNAs and hexose transporter genes of grape berry (*Vitis vinifera VvHT,* Y09590). The sequence alignment was performed using PILEUP from the Pacific Biosciences Research Center, University of Hawaii at Manoa, available at http://www.pbrc.hawaii.edu/ (March 26, 2007).

```
            1                                                    50
      CpHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       VvHT ATGCCGGCTG TCGGAGGCTT TGATAAGGGT ACCGGGAAGG CCTATCCCGG
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

            51                                                   100
      CpHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       VvHT TAACCTTACT CCTTACGTGA CTGTGACATG TGTTGTTGCA GCCATGGGTG
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

            101                                                  150
      CpHT1 ~~~~~~~~~~ ~GGATACGAT ATTGGGATCT CAGGGGGAGT GACGTCAATG
 F1080R2276 ~~~~~~~~~~ ~GGATACGAT ATTGGGATCT CAGGGGGAGT GACGTCAATG
 F1080R2322 ~~~~~~~~~~ ~GGATACGAT ATTGGGATCT CAGGGGGAGT GACGTCAATG
 F1080R2743 ~~~~~~~~~~ ~GGATACGAT ATTGGGATCT CAGGGGGAGT GACGTCAATG
 F1080R2622 ~~~~~~~~~~ ~GGATACGAT ATTGGGATCT CAGGGGGAGT GACGTCAATG
  F486R1076 ~~~~~~~~~~ ~GGATACGAT ATTGGGATCT CAGGGGGAGT GACGTCAATG
       VvHT GTTTGATCTT TGGTTACGAT ATTGGAATTT CTGGTGGGGT CACGTCCATG
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

            151                                                  200
      CpHT1 AACTCGTTTC TGAAGGAATT TTTCCCGGCG GTTTTCCGGA AAAAGGAAGA
 F1080R2276 AACTCGTTTC TGAAGGAATT TTTCCCGGCG GTTTTCCGGA AAAAGGAAGA
 F1080R2322 AACTCGTTTC TGAAGGAATT TTTCCCGGCG GTTTTCCGGA AAAAGGAAGA
 F1080R2743 AACTCGTTTC TGAAGGAATT TTTCCCGGCG GTTTTCCGGA AAAAGGAAGA
 F1080R2622 AACTCGTTTC TGAAGGAATT TTTCCCGGCG GTTTTCCGGA AAAAGGAAGA
  F486R1076 AACTCGTTTC TGAAGGAATT TTTCCCGGCG GTTTTCCGGA AAAAGGAAGA
       VvHT GCTCCGTTCT TGCAGAAGTT CTTCCCTTCT GTGTACCGGA AGGAGGCTTT
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.13. (continue)

```
          201                                                        250
     CpHT1 GGTATCGTCG ACTAACCAGT ACTGTCAGTA CGACAGTCCG ACACTTACGT
F1080R2276 GGTATCGTCG ACTAACCAGT ACTGTCAGTA CGACAGTCCG ACACTTACGT
F1080R2322 GGTATCGTCG ACTAACCAGT ACTGTCAGTA CGACAGTCCG ACACTTACGT
F1080R2743 GGTATCGTCG ACTAACCAGT ACTGTCAGTA CGACAGTCCG ACACTTACGT
F1080R2622 GGTATCGTCG ACTAACCAGT ACTGTCAGTA CGACAGTCCG ACACTTACGT
 F486R1076 GGTATCGTCG ACTAACCAGT ACTGTCAGTA CGACAGTCCG ACACTTACGT
      VvHT GGACAAGTCC ACGAATCAGT ACTGTAAGTT TGATAGTGAG ACACTAACGT
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          251                                                        300
     CpHT1 TGTTTACATC ATCGCTGTAT CTGGCGGCGC TTGTGGCGTC GCTGGTTGCG
F1080R2276 TGTTTACATC ATCGCTGTAT CTGGCGGCGC TTGTGGCGTC GCTGGTTGCG
F1080R2322 TGTTTACATC ATCGCTGTAT CTGGCGGCGC TTGTGGCGTC GCTGGTTGCG
F1080R2743 TGTTTACATC ATCGCTGTAT CTGGCGGCGC TTGTGGCGTC GCTGGTTGTG
F1080R2622 TGTTTACATC ATCGCTGTAT CTGGCGGCGC TTGTGGCGTC GCTGGTTGCG
 F486R1076 TGTTTACATC ATCGCTGTAT CTGGCGGCGC TTGTGGCGTC GCTGGTTGCG
      VvHT TGTTCACATC GTCGCTTTAT CTGGCTGCTC TTCTCTCGTC GCTGGTGGCC
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          301                                                        350
     CpHT1 GCGACGGTGA CAAGAAAGTT CGGTCGGAAA CTGTCGATGC TGTTTGGCGG
F1080R2276 GCGACGGTGA CAAGAAAGTT CGGTCGGAAA CTGTCGATGC TGTTTGGCGG
F1080R2322 GCGACGGTGA CAAGAAAGTT CGGTCGGAAA CTGTCGATGC TGTTTGGCGG
F1080R2743 GCGACGGTGA CAAGAAAGTT CGGTCGGAAA CTGTCGATGC TGTTTGGCGG
F1080R2622 GCGACGGTGA CAAGAAAGTT CGGTCGGAAA CTGTCGATGC TGTTTGGCGG
 F486R1076 GCGACGGTGA CAAGAAAGTT CGGTCGGAAA CTGTCGATGC TGTTTGGCGG
      VvHT GCGACGGTGA CCCGAAAGTT CGGGAGAAAG CTGTCAATGC TATTCGGAGG
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          351                                                        400
     CpHT1 CGTCCTGTTC TGCGCCGGTG CCATCATTAA TGGCTTCGCT AAAGCTGTTT
F1080R2276 CGTCCTGTTC TGCGCCGGTG CCATCATTAA TGGCTTCGCT AAAGCTGTTT
F1080R2322 CGTCCTGTTC TGCGCCGGTG CCATCATTAA TGGCTTCGCT AAAGCTGTTT
F1080R2743 CGTCCTGTTC TGCGCCGGTG CCATCATTAA TGGCTTCGCT AAAGCTGTTT
F1080R2622 CGTCCTGTTC TGCGCCGGTG CCATCATTAA TGGCTTCGCT AAAGCTGTTT
 F486R1076 CGTCCTGTTC TGCGCCGGTG CCATCATTAA TGGCTTCGCT AAAGCTGTTT
      VvHT ACTGCTCTTT TGTGCCGGTG CCATCATCAA TGGCGCTGCT AAAGCTGTTT
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          401                                                        450
     CpHT1 GGATGTTGAT TCTCGGCAGA ATTTTGTTGG GTTTTGGCAT CGGTTTTGCC
F1080R2276 GGATGTTGAT TCTCGGCAGA ATTTTGTTGG GTTTTGGCAT CGGTTTTGCC
F1080R2322 GGATGTTGAT TCTCGGCAGA ATTTTGTTGG GTTTTGGCAT CGGTTTTGCC
F1080R2743 GGATGTTGAT TCTCGGCAGA ATTTTGTTGG GTTTTGGCAT CGGTTTTGCC
F1080R2622 GGACGTTGAT TCTCGGCAGA ATTTTGTTGG GTTTTGGCAT CGGTTTTGCC
 F486R1076 GGATGTTGAT TCTCGGCAGA ATTTTGTTGG GTTTTGGCAT CGGTTTTGCC
      VvHT GGATGTTGAT TGTCGGTCGT ATACTGCTGG GTTTTGGTAT TGGGTTTGCC
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.13. (continue)

```
           451                                                        500
     CpHT1 AATCAGTCTG TACCACTCTA CCTCTCTGAG ATGGCTCCTT ACAGATATAG
F1080R2276 AATCAGTCTG TACCACTCTA CCTCTCTGAG ATGGCTCCTT ACAGATATAG
F1080R2322 AATCAGTCTG TACCACTCTA CCTCTCTGAG ATGGCTCCTT ACAGATATAG
F1080R2743 AATCAGTCTG TACCACTCTA CCTCTCTGAG ATGGCTCCTT ACAGATATAG
F1080R2622 AATCAGTCTG TACCACTCTA CCTCTCTGAG ATGGCTCCTT ACAGATATAG
 F486R1076 AATCAGTCTG TACCACTCTA CCTCTCTGAG ATGGCTCCTT ACAGATATAG
      VvHT AATCAGTCTG TGCCGCTCTA CCTCTCTGAG ATGGCTCCAT ACAAATACAG
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

           501                                                        550
     CpHT1 AGGAGCATTA AACATTGGAT TCCAATTGTC CATCACAATT GGTATTCTTG
F1080R2276 AGGAGCATTA AACATTGGAT TCCAATTGTC CATCACAATT GGTATTCTTG
F1080R2322 AGGAGCATTA AACATTGGAT TCCAATTGTC CATCACAATT GGTATTCTTG
F1080R2743 AGGAGCATTA AACATTGGAT TCCAATTGTC CATCACAATT GGTATTCTTG
F1080R2622 AGGAGCATTA AACATTGGAT TCCAATTGTC CATCACAATT GGTATTCTTG
 F486R1076 AGGAGCATTA AACATTGGAT TCCAATTGTC CATCACAATT GGTATTCTTG
      VvHT AGGAGCCCTC AACATTGGCT TCCAATTATC CATCACAATT GGTATTCTTG
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

           551                                                        600
     CpHT1 TTGCCAATGT ATTGAATTTC TTCTTTGCAA AAATCAAAGG AGGTTGGGGA
F1080R2276 TTGCCAATGT ATTGAATTTC TTCTTTGCAA AAATCAAAGG AGGTTGGGGA
F1080R2322 TTGCCAATGT ATTGAATTTC TTCTTTGCAA AAATCAAAGG AGGTTGGGGA
F1080R2743 TTGCCAATGT ATTGAATTTC TTCTTTGCAA AAATCAAAGG AGGTTGGGGA
F1080R2622 TTGCCAATGT ATTGAATTTC TTCTTTGCAA AAATCAAAGG AGGTTGGGGA
 F486R1076 TTGCCAATGT ATTGAATTTC TTCTTTGCAA AAATCAAAGG AGGTTGGGGA
      VvHT TGGCCAATAT ATTGAACTAC TTCTTTGCAA AGATCAAGGG GGGTTGGGGA
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

           601                                                        650
     CpHT1 TGGAGACTGA GCTTAGGAGG TGCAGTAGTT CCAGCTCTAA TCATCGCCAT
F1080R2276 TGGAGACTGA GCTTAGGAGG TGCAGTAGTT CCAGCTCTAA TCATCGCCAT
F1080R2322 TGGAGACTGA GCTTAGGAGG TGCAGTAGTT CCAGCTCTAA TCATCGCCAT
F1080R2743 TGGAGACTGA GCTTAGGAGG TGCAGTAGTT CCAGCTCTAA TCATCGCCAT
F1080R2622 TGGAGACTGA GCTTAGGAGG TGCAGTAGTT CCAGCTCTAA TCATCGCCAT
 F486R1076 TGGAGACTGA GCTTAGGAGG TGCAGTAGTT CCAGCTCTAA TCATCGCCAT
      VvHT TGGAGATTGA GCTTGGGTGG CGCTGTGGTC CCTGCGCTCA TCATCACCGT
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

           651                                                        700
     CpHT1 TGGATCGTTA ATCCTCCCCG ATACACCCAA CTCCATGATC GAACGAGGCC
F1080R2276 TGGATCGTTA ATCCTCCCCG ATACACCCAA CTCCATGATC GAACGAGGCC
F1080R2322 TGGATCGTTA ATCCTCCCCG ATACACCCAA CTCCATGATC GAACGAGGCC
F1080R2743 TGGATCGTTA ATCCTCCCCG ATACACCCAA CTCCATGATC GAACGAGGCC
F1080R2622 TGGATCGTTA ATCCTCCCCG ATACACCCAA CTCCATGATC GAACGAGGCC
 F486R1076 TGGATCGTTA ATCCTCCCCG ATACACCCAA CTCAATGATC GAACG~~~~~
      VvHT CGGGTCCCTT GTCCTCCCGG ACACACCCAA CTCCATGATC GAGCGTGGCC
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.13. (continue)

```
          701                                                       750
     CpHT1 AGGTAGAAGC AGCTAAAGAG AAATTAAGGA GAATTCGGGG TGTCAACAAC
F1080R2276 AGGTAGAAGC AGCTAAAGAG AAATTAAGGA GAATTCGGGG TGTCAACAAC
F1080R2322 AGGTAGAAGC AGCTAAAGAG AAATTAAGGA GAATTCGGGG TGTCAACAAC
F1080R2743 AGGTAGAAGC AGCTAAAGAG AAATTAAGGA GAATTCGGGG TGTCAACAAC
F1080R2622 AGGTAGAAGC AGCTAAAGAG AAATTAAGGA GAATTCGGGG TGTCAACAAC
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT AGCACGAGGG AGCGAAAACA AAACTGAGAA GAATCCGGGG TGTCGATGAT
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          751                                                       800
     CpHT1 GTGGACGAAG AGTTAAAAGA CTTAGTTGCA GCAAGTGAAG CTTCGAAATT
F1080R2276 GTGGACGAAG AGTTAAAAGA CTTAGTTGCA GCAAGTGAAG CTTCGAAATT
F1080R2322 GTGGACGAAG AGTTAAAAGA CTTAGTTGCA GCAAGTGAAG CTTCGAAATT
F1080R2743 GTGGACGAAG AGTTAAAAGA CTTAGTTGCA GCAAGTGAAG CTTCGAAATT
F1080R2622 GTGGACGAAG AGTTAAAAGA CTTAGTTGCA GCAAGTGAAG CTTCGAAATT
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT GTTGAAGAGG AATTCAATGA CCTTGTTGTA GCCAGTGAGG CCTCCAAGCT
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~TT

          801                                                       850
     CpHT1 GGTAGAACAT CCATGGAGAA ACTTGTTACA AAGAAAATAC AGGCCTCATC
F1080R2276 GGTAGAACAT CCATGGAGAA ACTTGTTACA AAGAAAATAC AGGCCTCATC
F1080R2322 GGTAGAACAT CCATGGAGAA ACTTGTTACA AAGAAAATAC AGGCCTCATC
F1080R2743 GGTAGAACAT CCATGGAGAA ACTTGTTACA AAGAAAATAC AGGCCTCATC
F1080R2622 GGTAGAACAT CCATGGAGAA ACTTGTTACA AAGAAAATAC AGGCCTCATC
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT TGTTGAGCAC CCCTGGAGAA ATCTCTTGCA GAGGAAGTAC AGGCCACACC
  GR3F1645 GGTAGAACAT CCATGGAGAA ACTTGTTACA AAGAAAATAC AGGCCTCATC

          851                                                       900
     CpHT1 TCACCATGGC TATCATGATC CCATTCTTCC AGCAGCTAAC TGGAATTAAT
F1080R2276 TCACCATGGC TATCATGATC CCATTCTTCC AGCAGCTAAC TGGAATTAAT
F1080R2322 TCACCATGGC TATCATGATC CCATTCTTCC AGCAGCTAAC TGGAATTAAT
F1080R2743 TCACCATGGC TATCATGATC CCATTCTTCC AGCAGCTAAC TGGAATTAAT
F1080R2622 TCACCATGGC TATCATGATC CCATTCTTCC AGCAGCTAAC TGGAATTAAT
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT TCACAATGGC CATCCTCATT CCCTTCTTCC AGCAGCTTAC CGGGATTAAT
  GR3F1645 TCACCATGGC TATCATGATC CCATTCTTCC AGCAGCTAAC TGGAATTAAT

          901                                                       950
     CpHT1 GTCATCATGT TTTATGCTCC TGTTTTATTC AACACAATTG GGTTTGGCAG
F1080R2276 GTCATCATGT TTTATGCTCC TGTTTTATTC AACACAATTG GGTTTGGCAG
F1080R2322 GTCATCATGT TTTATGCTCC TGTTTTATTC AACACAATTG GGTTTGGCAG
F1080R2743 GTCATCATGT TTTATGCTCC TGTTTTATTC AACACAATTG GGTTTGGCAG
F1080R2622 GTCATCATGT TTTATGCTCC TG~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT GTCATTATGT TTTATGCCCC TGTTCTCTTC AAAACTATTG GCTTTGCGGA
  GR3F1645 GTCATCATGT TTTATGCTCC TGTTTTATTC AACACAATTG GGTTTGGCAG
```

Figure 5.13. (continue)

```
          951                                                  1000
     CpHT1 TGACGCCTCT CTCATGTCTG CTGTAATTAC CGGAATTGTA AATGTCGGTG
F1080R2276 TGACGCCTCT CTCATGTCTG CTGTAATTAC CGGAATTGTA AATGTCGGTG
F1080R2322 TGACGCCTCT CTCATGTCTG CTGTAATTAC CGGAATTGTA AATGTCGGTG
F1080R2743 TGACGCCTCT CTCATGTCTG CTGTAATTAC CGGAATTGTA AATGTCGGTG
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT TGATGCTTCC CTGATGTCTG CTGTGATAAC CGGGCGGGTT AATGTTCTTG
  GR3F1645 TGACGCCTCT CTCATGTCTG CTGTAATTAC CGGAATTGTA AATGTCGGTG

          1001                                                 1050
     CpHT1 CAACTTTGGT TTCAATCTAT GGAGTCGATA AATGGGGAAG ACGATTCCTT
F1080R2276 CAACTTTGGT TTCAATCTAT GGAGTCGATA AATGGGGAAG ACGATTCCTT
F1080R2322 CAACTTTGGT TTCAATCTAT GGAGTCGATA AATGGGGAAG ACGATTCCTT
F1080R2743 CAACTTTGGT TTCAATCTAT GGAGTC~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT CAACCATAGT TTCAATCTAC GGTGTTGATA AGTGGGTAAG AAGGTTTCTT
  GR3F1645 CAACTTTGGT TTCAATCTAT GGAGTCGATA AATGGGGAAG ACGATTCCTT

          1051                                                 1100
     CpHT1 TTCCTCGAGG GAGGAGTTCA AATGTTAATA TGCCAGATTG TGGTAGCAGC
F1080R2276 TTCCTCGAGG GAGGAGTTC~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 TTCCTCGAGG GAGGAGTTCA AATGTTAATA TGCCAGATTG TGGTAGCAGC
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT TTCCTTGAGG GTGGCACTCA AATGCTCATA TGTCAGGTTA TTGTGGCAAC
  GR3F1645 TTCCTCGAGG GAGGAGTTCA AATGTTAATA TGCCAGATTG TGGTAGCAGC

          1101                                                 1150
     CpHT1 CTCCATTGGA GCTAAATTTG GGATCAATGG CAACCCTGGA GATTTACCAA
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 CTCCATTGGA G~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT GTGCATTTTG GTTAAATTCG GAGTGGATGG AGAACCTTGG TGCTTGCCCA
  GR3F1645 CTCCATTGGA GCTAAATTTG GGATCAATGG CAACCCTGGA GATTTACCAA

          1151                                                 1200
     CpHT1 AATGGTATGC AATTGTAGTG GTGCTATTCA TCTGTATTTA CGTGGCCGGA
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT AGTGGTATGC CATAGTTGTG GTGCTGTTCA TTTGCGTCTA TGTTTCAGGG
  GR3F1645 AATGGTATGC AATTGTAGTG GTGCTATTCA TCTGTATTTA CGTGGCCGGA
```

Figure 5.13. (continue)

```
          1201                                                  1250
     CpHT1 TTTGCGTGGT CTTGGGGGCC TCTCGGGTGG CTCGTGCCGA GTGAAATCTT
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT TTTGCATGGT CCTGGGGACC TCTAGGTTGG TTGGTCCCTA GTGAAATTTT
  GR3F1645 TTTGCGTGGT CTTGGGGGCC TCTCGGGTGG CTCGTGCCGA GTGAAATCTT

          1251                                                  1300
     CpHT1 CCCTCTTGAA ATCAGATCAG CAGCTCAGAG TATCAATGTG TCGGTGAATA
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT CCCCCTGGAA ATCCGATCTG CTGCACAGAG TGTAAACGTC TCCGTTAACA
  GR3F1645 CCCTCTTGAA ATCAGATCAG CAGCTCAGAG TATCAATGT~ ~~~~~~~~~~

          1301                                                  1350
     CpHT1 TGTTCTTCAC ATTTATAGTG GCACAAATAT TTTTGACAAT GCTTTGTCAT
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT TGTTTTTCAC ATTCATCATA GCCCAAATCT TCTTAAATAT GCTGTGTCAC
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1351                                                  1400
     CpHT1 TTGAAGTTTG GACTGTTCAT TTTCTTTGCC TTCTTTGTGG TTATCATGTC
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT ATGAAGTTTG GTTTGTTCCT CTCCTTTGCC TTCTTTGTGG TGGTGATGTC
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1401                                                  1450
     CpHT1 AATCTTCATC TACTATTTCT TGCCGGAGAC AAAGGGAATC CCCATTGAAG
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT CTTCTTCATT TACTTCTTCT TGCCTGAGAC CAAAGGCATC CCAATTGAAG
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.13. (continue)

```
          1451                                                      1500
     CpHT1 AAATGAGTAA GGTCTGGAAG TCTCACTGGT TCTGGTCCAG GTTTGTGGAA
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT AGATGGCTGA AGTATGGAAA AGTCACTGGT TCTGGTCCCG GTATGTCAAC
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1501                                                      1550
     CpHT1 GATGATGGTT ATGGCCATCA TGGAAATCTT GAGATGGGAA AAGGCAACCT
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT GATGGTTCTT ACAGCGGCGT CGAACTGGTC AAGGAAAACT ACCCTGTTAA
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1551                                                      1600
     CpHT1 GGGACCAAAG AATGTGTGAT TCATTTTTCT TTTTTTTTTT TTTTGTTTTT
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT GAATGTATGA ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1601                                                      1650
     CpHT1 TTCTGGGTTT TAAAGAATCG GAAGCAGTAA ATAGACTTTT TAGTTTGTTT
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1651                                                      1700
     CpHT1 GTATTACGGA ATTACGAAAA ACTCTCGATA ATACAGATCA ATTCTATTCA
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.13. (continue)

```
          1701                                                      1750
     CpHT1 ATAATCATAG TAAGTTATGT TCTAAAAAAA AAAAAAAAAA AAAAAAAAAA
F1080R2276 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2322 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2743 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
F1080R2622 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
 F486R1076 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      VvHT ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          1751
     CpHT1 AAAA
F1080R2276 ~~~~
F1080R2322 ~~~~
F1080R2743 ~~~~
F1080R2622 ~~~~
 F486R1076 ~~~~
      VvHT ~~~~
  GR3F1645 ~~~~
```

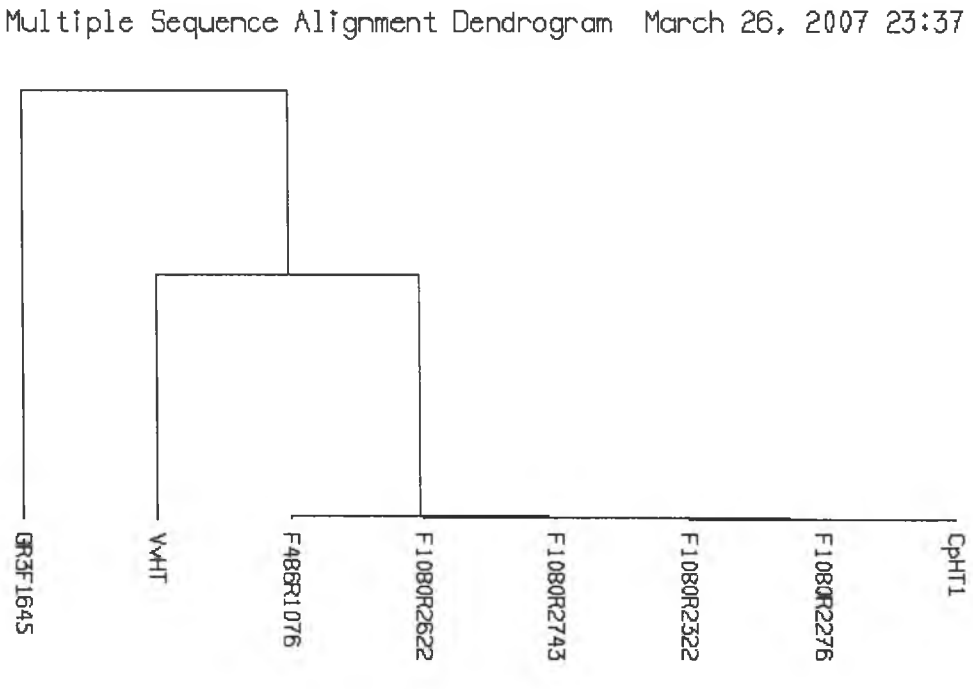

Figure 5.13. (continue)

Figure 5.14. The multiple comparisons between the partial *CpHT1* cDNAs and two clones amplified from GeneRacer 3' primer indicated that the same gene was amplified. The sequence alignment was performed using PILEUP from the Pacific Biosciences Research Center, University of Hawaii at Manoa, available at http://www.pbrc.hawaii.edu/ (March 26, 2007).

```
        1                                                   50
   CpHT1 GGATACGATA TTGGGATCTC AGGGGGAGTG ACGTCAATGA ACTCGTTTCT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        51                                                  100
   CpHT1 GAAGGAATTT TTCCCGGCGG TTTTCCGGAA AAAGGAAGAG GTATCGTCGA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        101                                                 150
   CpHT1 CTAACCAGTA CTGTCAGTAC GACAGTCCGA CACTTACGTT GTTTACATCA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        151                                                 200
   CpHT1 TCGCTGTATC TGGCGGCGCT TGTGGCGTCG CTGGTTGCGG CGACGGTGAC
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        201                                                 250
   CpHT1 AAGAAAGTTC GGTCGGAAAC TGTCGATGCT GTTTGGCGGC GTCCTGTTCT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        251                                                 300
   CpHT1 GCGCCGGTGC CATCATTAAT GGCTTCGCTA AAGCTGTTTG GATGTTGATT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        301                                                 350
   CpHT1 CTCGGCAGAA TTTTGTTGGG TTTTGGCATC GGTTTTGCCA ATCAGTCTGT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        351                                                 400
   CpHT1 ACCACTCTAC CTCTCTGAGA TGGCTCCTTA CAGATATAGA GGAGCATTAA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        401                                                 450
   CpHT1 ACATTGGATT CCAATTGTCC ATCACAATTG GTATTCTTGT TGCCAATGTA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

```
          451                                                        500
   CpHT1 TTGAATTTCT TCTTTGCAAA AATCAAAGGA GGTTGGGGAT GGAGACTGAG
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          501                                                        550
   CpHT1 CTTAGGAGGT GCAGTAGTTC CAGCTCTAAT CATCGCCATT GGATCGTTAA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          551                                                        600
   CpHT1 TCCTCCCCGA TACACCCAAC TCCATGATCG AACGAGGCCA GGTAGAAGCA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          601                                                        650
   CpHT1 GCTAAAGAGA AATTAAGGAG AATTCGGGGT GTCAACAACG TGGACGAAGA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          651                                                        700
   CpHT1 GTTAAAAGAC TTAGTTGCAG CAAGTGAAGC TTCGAAATTG GTAGAACATC
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~TTG GTAGAACATC
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          701                                                        750
   CpHT1 CATGGAGAAA CTTGTTACAA AGAAAATACA GGCCTCATCT CACCATGGCT
GR3F1645 CATGGAGAAA CTTGTTACAA AGAAAATACA GGCCTCATCT CACCATGGCT
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          751                                                        800
   CpHT1 ATCATGATCC CATTCTTCCA GCAGCTAACT GGAATTAATG TCATCATGTT
GR3F1645 ATCATGATCC CATTCTTCCA GCAGCTAACT GGAATTAATG TCATCATGTT
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          801                                                        850
   CpHT1 TTATGCTCCT GTTTTATTCA ACACAATTGG GTTTGGCAGT GACGCCTCTC
GR3F1645 TTATGCTCCT GTTTTATTCA ACACAATTGG GTTTGGCAGT GACGCCTCTC
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          851                                                        900
   CpHT1 TCATGTCTGC TGTAATTACC GGAATTGTAA ATGTCGGTGC AACTTTGGTT
GR3F1645 TCATGTCTGC TGTAATTACC GGAATTGTAA ATGTCGGTGC AACTTTGGTT
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

          901                                                        950
   CpHT1 TCAATCTATG GAGTCGATAA ATGGGGAAGA CGATTCCTTT TCCTCGAGGG
GR3F1645 TCAATCTATG GAGTCGATAA ATGGGGAAGA CGATTCCTTT TCCTCGAGGG
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
```

Figure 5.14. (continue)

```
         951                                                       1000
   CpHT1 AGGAGTTCAA ATGTTAATAT GCCAGATTGT GGTAGCAGCC TCCATTGGAG
GR3F1645 AGGAGTTCAA ATGTTAATAT GCCAGATTGT GGTAGCAGCC TCCATTGGAG
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

        1001                                                      1050
   CpHT1 CTAAATTTGG GATCAATGGC AACCCTGGAG ATTTACCAAA ATGGTATGCA
GR3F1645 CTAAATTTGG GATCAATGGC AACCCTGGAG ATTTACCAAA ATGGTATGCA
GR3F2259 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ATTTACCAAA ATGGTATGCA

         1051                                                      1100
   CpHT1 ATTGTAGTGG TGCTATTCAT CTGTATTTAC GTGGCCGGAT TTGCGTGGTC
GR3F1645 ATTGTAGTGG TGCTATTCAT CTGTATTTAC GTGGCCGGAT TTGCGTGGTC
GR3F2259 ATTGTAGTGG TGCTATTCAT CTGTATTTAC GTGGCCGGAT TTGCGTGGTC

         1101                                                      1150
   CpHT1 TTGGGGGCCT CTCGGGTGGC TCGTGCCGAG TGAAATCTTC CCTCTTGAAA
GR3F1645 TTGGGGGCCT CTCGGGTGGC TCGTGCCGAG TGAAATCTTC CCTCTTGAAA
GR3F2259 TTGGGGGCCT CTCGGGTGGC TCGTGCCGAG TGAAATCTTC CCTCTTGAAA

         1151                                                      1200
   CpHT1 TCAGATCAGC AGCTCAGAGT ATCAATGTGT CGGTGAATAT GTTCTTCACA
GR3F1645 TCAGATCAGC AGCTCAGAGT ATCAATGT-- ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 TCAGATCAGC AGCTCAGAGT ATCAATGTGT CGGTGAATAT GTTCTTCACA

         1201                                                      1250
   CpHT1 TTTATAGTGG CACAAATATT TTTGACAATG CTTTGTCATT TGAAGTTTGG
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 TTTATAGTGG CACAAATATT TTTGACAATG CTTTGTCATT TGAAGTTTGG

         1251                                                      1300
   CpHT1 ACTGTTCATT TTCTTTGCCT TCTTTGTGGT TATCATGTCA ATCTTCATCT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ACTGTTCATT TTCTTTGCCT TCTTTGTGGT TATCATGTCA ATCTTCATCT

         1301                                                      1350
   CpHT1 ACTATTTCTT GCCGGAGACA AAGGGAATCC CCATTGAAGA AATGAGTAAG
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ACTATTTCTT GCCGGAGACA AAGGGAATCC CCATTGAAGA AATGAGTAAG

         1351                                                      1400
   CpHT1 GTCTGGAAGT CTCACTGGTT CTGGTCCAGG TTTGTGGAAG ATGATGGTTA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 GTCTGGAAGT CTCACTGGTT CTGGTCCAGG TTTGTGGAAG ATGATGGTTA

         1401                                                      1450
   CpHT1 TGGCCATCAT GGAAATCTTG AGATGGGAAA AGGCAACCTG GGACCAAAGA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 TGGCCATCAT GGAAATCTTG AGATGGGAAA AGGCAACCTG GGACCAAAGA
```

Figure 5.14. (continue)

```
         1451                                                      1500
   CpHT1 ATGTGTGATT CATTTTTCTT TTTTTTTTTT TTTGTTTTTT TCTGGGTTTT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 ATGTGTGATT CATTTTTCTT TTTTTTTTTT TTTGTTTTTT TCTGGGTTTT

         1501                                                      1550
   CpHT1 AAAGAATCGG AAGCAGTAAA TAGACTTTTT AGTTTGTTTG TATTACGGAA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 AAAGAATCGG AAGCAGTAAA TAGACTTTTT AGTTTGTTTG TATTACGGAA

         1551                                                      1600
   CpHT1 TTACGAAAAA CTCTCGATAA TACAGATCAA TTCTATTCAA TAATCATAGT
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
GR3F2259 TTACGAAAAA CTCTCGATAA TACAGATCAA TTCTATTCAA TAATCATAGT

         1601                                                1643
   CpHT1 AAGTTATGTT CTAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAA
GR3F1645 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~
GR3F2259 AAGTTATGTT CTAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA~
```

Figure 5.14. (continue)

Figure 5.15. *CpHT1* cDNA was homologous with the sequence in the papaya genome supercontig_1226 (http://cgpbr.hawaii.edu/papaya/) . Supercontig_1226 contained a similar motif to the hexose transporter gene of Poplar tMST2.2 (AJ698938). Four potential exons were displayed in capitol letters between three introns, displayed in small letter. The translation initiation codon (ATG) was predicted at base 284 in the red box. The size and location of the 5'UTR were unclear. A possible CCAAT motif (double underline) was located at – 211 nt upstream from the adenine residue of the translational start codon (ATG). Three possible TATA motifs (underline) were located at – 199, – 119 and – 107 nt. The initiator site (Inr), pyrimidine rich region was shown in small italic letters. The possible transcription termination signal, AAT(U)AAA displayed in bold capitol letters. *CpHT1* contained a short sequence of poly T (U) tract (underline). The stop codon (TGA) was predicted at base 3,391 (violet box). The letter 'T' in the green box showed the poly A site. CAG splice site showed in gray boxes.

attaagatattacttttgacagctcaacagttgatgatcccagttgagaaatgggtaatgggtacagtac
a**CCAAT**ctccttcTATATATAcacttgccatgcctgtgctaa*cctctatcttctg*cattaaa*actctcgtct*
*tcttcatttcctccctctg*catTATATATATAcaTATAATATA*tcttccattttg*ccacatatagttg*tttt*
*tct*gctag*ccttcctttttgattttcccactgccc*aaagagaagaaagaaagaaagaaagaaaaaaaaaa
CCTGCACCAGGAGGAATTGCGCCGGCTGAGCCCGGCAGGGAATACCCCGGTAATCTTACCCCATTCGTC
ACTGTAACATGTATCGTCGCCGCCATGGGTGGACTGATCTTTGGCTACGATATTGGGATCTCAGgtatata
catatatgaacatattgttatatatcgcttgctatatttcatgcatttttgttgatgaaaatggttttga
tggttgttgaccgttgactaaaccagGGGGAGTGACGTCAATGAACTCGTTTCTGAAGGAATTTTTCCCGG
CGGTTTTCCGGAAAAAGGAAGAAGTATCGTCGACTAACCAGTACTGTCAGTACGACAGTCCGACACTTACG
TTGTTTACATCATCGCTGTATCTGGCGGCGCTTGTGGCGTCGCTGGTTGCGGCGACGGTGACAAGAAAGTT
CGGTCGGAAACTGTCGATGCTGTTTGGGCGGCGTCCTGTTCTGCGCCGGTGCCATCATTAATGGCTTCGCT
AAAGCTGTTTGGATGTTGATTCTCGGCAGAATTTTGTTGGGTTTTGGCATCGGTTTTGCCAATCAGgtatt
ctttgttttaattctagtaatttgcaaaatttgttaacacaagtttttttatttttggtttttttttgggt
tcggagaataaaacattacatatgaattttggtgtcagattttgtctgtttggtgtttttgttctctgag
aaactccttttttcctcagatctggtttctgtaacgccttttttttttttaactttacaaagtgacaacttg
atacttttgtcccacctttttgatttcctcttctgattatgatatgggttccctgttcatttgcaagttt
tgaaaagattagtctttgatttatgcttttggagatctgaaagtaaggataatgatgataatattgttata
ataatgttttgtttggctagcgagaaaagggaggaaaggaaaggaaaggaaaggaaatccatcaagttttt
tttctttaaaaaaaaaaaaaaaaactccatttttctcaacagccaaacaaatgggtgttaaaacttaattcc
ttttttttaatttccgagttggagaaagggaaattttgtgtctctctgcaattatataattgcatagaaaca
aaaagaggcaaagaataatttgagatggagaacatgatagagttcttggtttatttatattacgtatatta
catatacacacatgtatatgtatatatatttatttacaagtaatatatggcctgtttttctatccatgtga
agatccttttgttttcttctgacttttttactcccttgatgaaattctccgttgaataaagaatagaaaac
tctgataccagctttgaataattcaaagatgatgaaactaaaaaaccagtccttatattattgaatctttt
ttttatatatataataatttttcaaagcttgattctttcattaaaaacccagtttttttttcttcttcctct
tgtttttcactgatgaacccaggttagtactgactacttgacaacaacatgttgtggtccctgaggaagac
atttttttctatatttagtttatcctaatcttgggtaatttttttttttttttttacagTCTGTACCACTCT
ACCTCTCTCTGAGATGGCTCCTTACAGATATAGAGGAGCATTAAACATTGGATTCCAATTGTCCATCACAATT
GGTATTCTTGTTGCCAATGTATTGAATTTCTTCTTTGCAAAAATCAAAGGAGGTTGGGGATGGAGACTGAG
CTTAGGAGGTGCAGTAGTTCCAGCTCTAATCATCGCCATTGGATCGTTAATCCTCCCCGATACACCCAACT
CCATGATCGAACGAGGCCAGGTAGAAGCAGCTAAAGAGAAATTAAGGAGAATTCGGGGTGTCAACAACGTG
GACGAAGAGTTAAAAGACTTAGTTGCAGCAAGTGAAGCTTCGAAATTGGTAGAACATCCATGGAGAAACTT
GTTACAAAGAAAATACAGGCCTCATCTCACCATGGCTATCATGATCCCATTCTTCCAGCAGCTAACTGGAA
TTAATGTCATCATGTTTTATGCTCCTGTTTTATTCAACACAATTGGGTTTGGCAGTGACGCCTCTCTCATG
TCTGCTGTAATTACCGGAATTGTAAATGTCGGTGCAACTTTGGTTTCAATCTATGGAGTCGATAAATGGGG
AAGACGATTCCTTTTCCTCGAGGGAGGAGTTCAAATGTTAATATGCCAGgtttgaattactcaaattcgtg
tctttttaatcctatctttttttctattgatgagtatcgagtaaagtcttgtgcaaaattcggtatgcagAT
TGTGGTAGCAGCCTCCATTGGAGCTAAATTTGGGATCAATGGCAACCCTGGAGATTTACCAAAATGGTATG
CAATTGTAGTGGTGCTATTCATCTGTATTTACGTGGCCGGATTTGCGTGGTCTTGGGGGCCTCTCGGGTGG
CTCGTGCCGAGTGAAATCTTCCCTCTTGAAATCAGATCAGCAGCTCAGAGTATCAATGTGTCGGTGAATAT
GTTCTTCACATTTATAGTGGCACAAATATTTTTGACAATGCTTTGTCATTTGAAGTTTGGACTGTTCATTT

TCTTTGCCTTCTTTGTGGTTATCATGTCAATCTTCATCTACTATTTCTTGCCGGAGACAAAGGGAATCCCC
ATTGAAGAAATGAGTAAGGTCTGGAAGTCTCACTGGTTCTGGTCCAGGTTTGTGGAAGATGATGGTTATGG
CCATCATGGAAATCTTGAGATGGGAAAAGGCAACCTGGGACCAAAGAATGTG[redacted]TTCATTTTTCTTTTTT
TTTTTTTTTGTTTTTTTCTGGGTTTTAAAGAATCGGAAGCAGTAAATAGACTTTTTAGTTTGTTTGTATTA
CGGAATTACGAAAAACTCTCGATAATACAGATCAATTCTATTC**AATAA**TCATAGTAAGTTATGTTC[redacted]tatt
atctcttaaatttcgattttcgttatcggtttattaaaagattatatattataatagatgtgtgtatgtat
aaactaatttcatcactgaacttcaaaaatcaaactatagagaaatagagggtcatttgtggaattatatc
taagatgaatttggttcatgcacctacaaaaaaacaaaaattaattaaatggttaaattaataagacatcc
gaacgtgggtttcctgtttggtaggggcttgtatgagaagaaagggctgttgggcatcatatagaaatccg
aataggaaaccaaaaaaaaaaaaaaatgcaaaagaggaaacatgaattttgaaggatatgaaatgaccgaa
gaaataatatgattaattatgtattactttcacttttcgcttttttattttcctcaaacactttggaatttt
aaaataataataatatttttttttaaaaggtttttaaactcaaaaatatatatatatatatatatattaagt
gttttttaattagaaaaaaagacataattttgtggtaaagtgacaatcttaacgaacttgattaccctaca
ttcgcataaa**AATA**AATAAAaaact**AATAAA**attaatattagagagtagatgtgacattatttatgat**AAT**
AAActttaaacgaatataaaatatataaattctgat**AATAAA**aaataatttttaaattaatttttacttacca
at**AATAAA**gttataat**AATAAA**aagaattattatattatcttacaactcatttaataataataataatttt
ttaaataatactataaaatagttaagaaaatatatcaaatactgacatataataataactataataattta
cttgtttaatatttatttatttaattatttcacaaaaaaaaatatatgatgtctacatagtttaaatttttaa
ataatttccaaaatttttaaaacatcttaaattaaaaacaatatggatgcatacttttttattaatcggtcta
ctttctaactcagaaacttgcattaccaacttaaacaaatttaatttacataaattcacaagaacagaaat
atgatttattattaaaaattcattaaataatcaaaattctttattaggatattgtaaaattgaagaagcat
tttcatttcggacaaa**AATAAA**aaatatgataatgatatatatatatatatatatatatatatgtaggggt
cacgttgggttaatgaaaattttcattgtttaggttggttaattataagataatattaattaatcatatct
aacccattgttaatttcttaattagcagattacttgctaacctttaatttttcttgttttgattaatgggag
gatcagttgagagatttttttaactttataaaaaagaacattttatttttctcaattaatcttaaatctta
ataatttaatttgcagacgtataatttttatgtattcgatgtataattaagttaaattaacttatactttca
tccatacaatgagatatgaaattatattttcatatattaat**AATAAA**aaaaattacataatatcaatttat
ttaaaattaaa**AATAAA**atattataatttatttaatattatctttttgtgattctaaagtaattattaatt
tgtatcacttacgctattctgaattttcttttactagaatatgcttcttctttttttttaatttctttttt
ttttttctgaattctcaagaacatgaatggaaccgaacccatcaattattcaaccaaaagaagatgtgaaa
atttataataatgaa**AATAAAATAAA**aaaaatcaaactccaaatatgattaattaacagatgatatatttt
gattttcaaaactaaattagtcaataataattatatatatagcttttattatagagtaattatatatatttt
tgcatgtaaagtttgaagccatgatgtttatccatgtaagaatgactcagcaaatgaaattaaaacgtgtt
ataactcagaattttgatacaaaatacaataattatatggctcataatttgcataacataattgatttca
tttcattgttgattatatacatgcatatacacatatacatacacccacaataactcatactaggctcta
gagatatacatattatataataattacccttttatttcgaatatgctaaaagcatcgtgtgtgtgtgtgtgt
gtgtgtgtgtctatacatatatatatacagaaacatgaggtttgtaaattacccataaacgagaaagggga
aacaaaaaggagaaattaatgtgtggataaaaggtgtggaaaatcacatgactttatatatatatatatat
atatatatatatatatataaaagttgtagaagagaattaatccacgtattatacaccaatttatgtaataatt
aagaactagaagctttcggtggactcaaattaaacctcacattattcatatttccttaaagataagaattg
tatatatatatatatacatgtgtacacgtatgtgtatacatatagtggaaatatataaagagtttggtcac
ggtgttagaacgaagaaaaggaaaggagaatcgcatgagccgtggtttatgattatgggctacattaattc
actacaccatgcattctttctatatatatatatatatatatatattgttgttatgttatcatatatata
tataagtatgtatgtgataggtgtgtggaatttatttgttgggttacttgtcccacatatcttgattttat
tttatttatatcttttggattggaaaaaggataatataaaaat**AATAAA**aacaatttaattggtatcata
taatatatgcttttgatattttatcttctttcaacttttgttacgaaccacatcatgtttcatgcacctt

Figure 5.15. (continue)

ctcccccactttcctcactctctggatcttcttctttttttttctttttttttttaatttgaaaaagttata
gtttatatatatatataaatacacatatatatacacatctatattactaa**AATAAATAAA**attctaacagt
aatatatttttacccattaaattaattataaatttggtttaatttttat**AATAAA**gtataaattagagtaa
ttaaattaatatccctttaaattcttgtaagaaaaaagtgagacaaatagcatttgataaataatgaagct
ttttcccccctcaaaatgacaaatatatatagtaaattagcttttatatatacatacatagatatataagg
taaattgatgctttgctttaaatactagctttcatattatgtgtgtgtgtatatatatatatatgtacaca
agattttattctgaaaataataccaacaaggttggatcggaattgccattgctagaaaccatgtaggcaca
cacacacacacacacatatatatatatatatatatctttattgactttcattatcgcaggaacaaagttaagt
tgaaacatccagtaagaaaaaaattacaattaagatgtgggtcacaattgaatttttatttgtgtcaatat
aattggagataattattaattatctcaaattattttgtaatatgtatttgtatgtatgtatgtatatatat
aatatcagtttgtagagccagaaacacagatccattgacccacaaagagagcagtgcaattgagattaggg
ttgattgaatcaaaagtctgagtggtgagattgaactgctgggaaattccaatgttggctgggctttgcct
ttaaaacatcaactgcattcattgacccataattaattacttaattacataattaattaatgttaattacc
ctcatataacaactaatctttttttaactgaaccattcaaatcatgggtgtgcatacaatgatgataattaa
ataaggaaatcaacaataatcaaaaccaaatctgtccatttacattgttccacatgtgtttttatacgtac
atatctatatatatatatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnn
nnnnnnnnnnnnnnntatggaattttaccctcaatagtagtatgacatctagaacatttatgattcatttc
aaatgaa**AATAAA**

Figure 5.15. (continue)

Intron	5' splice site	3' splice site
1	GATCTCAGGTATA	TAAACCAGGGGGA
2	CCAATCAGGTATT	TTTTACAGTCTGT
3	TATGCCAGGTTTG	GTATGCAGATTGT
Conserve	BMWVYCAGGTWTD	KWWWVCAGDBKGW

Figure 5.16. The conserved regions (underline) of the 5' and 3' recognition splice sites (gray box).

Yellow box shows exonic sequences.

Figure 5.17. The multiple comparison and phylogenic tree for hexose transporter genes between the full length *CpHT1* cDNA and barrel medic *MtST1* (*Medicago truncatula,* U38651), *Arabidopsis thaliana AtSTP1* (X55350), tobacco *NtMST1* (*Nicotiana tabacum,* X66856), fava bean *VfSTP1* (*Vicia faba,* Z93775), castor bean *RcSCP* (*Ricinus communis,* L08196), poplar tMST2.2 (Populus tremula x Populus tremuloides, AJ698938), grape berry *VvHT* (*Vitis vinifera,* Y09590), peach fruit *PpSTP1* (*Prunus persica*, AF367454) and *PpSTP2* (AF367455) and tomato fruit *LeHT1* (*Lycopersicon esculentum,* AJ132223), *LeHT2* (AJ132224) and *LeHT3* (AJ132225). The sequence alignment was performed using PILEUP from the Pacific Biosciences Research Center, University of Hawaii at Manoa, available at http://www.pbrc.hawaii.edu/ (March 27, 2007).

```
          1                                                  50
    LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   NtMST1 ATTTCTACAG TTTACGATTA CAGTTCTCTC TCAACTTTTT TCTTTTTTCT
    MtST1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   VfSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    RcSCP ~~~~~~~~~~ ~CTCTGTTTA AGCTTCTTGT TCTATTTCTG TTTCTCTGCA
     VvHT ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    tMST2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Full-CpHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   AtSTP1 ~~~~~~~~~~ ~~~~~~~~~A AAAACCCATC CCATCAAAAA TAAACAAGAG
    LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    LeHT2 ~~~~~~~~~~ ~~~ATTTTAT TTTTCTTATA TAAAAAACAT ATCATCTTCT

          51                                                 100
    LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   NtMST1 TGATCATTTC TTGTGCCACA AGTTGTTCTT GTTTTGTGAT CAGCTCAGAA
    MtST1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~CTTCTT CTTGCTACTG AGGTCAGAAA
   VfSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~CTTCA TCGGAGAAAA
   PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    RcSCP AGGAAGTGTA CGTAGCATAT AACTAAGAGC CAAAAAAGAA AAAAAAATGC
     VvHT ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    tMST2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~GCTTGCAAA GAACGAGCAA
Full-CpHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~A
   AtSTP1 GGCCTAAAGA AGAATCCTAA AGACTTTACG GGTCTTGTTT AGGATAAAAG
    LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    LeHT2 TATTGTTATT TTAACTACCA CATTTGAAGC CGGAAAACTT AGTAACACGA
```

```
          101                                                       150
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 AGATGGCTGG TGGTGGTGGT ATTGGTCCCG GCAACGGGAA AGAATATCCC
     MtST1 AAATGGCTGG TGGTGGAATT CCCATTGGAG GGGGTAACAA AGAGTACCCC
    VfSTP1 AAATGCCTGC AGCCGGAATC CCCATCGGAG CGGGGAACAA GGAGTACCCC
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP CTGCAGTAGG AGGTATACCG CCCTCTGGTG GCAACAGGAA AGTGTACCCG
      VvHT ~~ATGCCGGC TGTCGGAGGC TTTGATAAGG GTACCGGGAA GGCCTATCCC
     tMST2 AAATGCCTGC AGTAGGGATA GCTGTCGGTG ACAACAAAAA GGAGTATCCG
Full-CpHT1 TGCCTGCACC AGGAGGAATT GCGCCGGCTG AGCCCGGCAG GGAATACCCC
    AtSTP1 AAATGCCTGC CGGTGGATTC GTCGTCGGGG ATGGCCAAAA GGCTTATCCC
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 TGGCCGGTGG AGGATTTACG ACCTCCGGTA ACGGAGGCAC GCATTTCGAG

          151                                                       200
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 GGCAATTTAA CTCTTTATGT TACCGTTACG TGCATTGTCG CTGCCATGGG
     MtST1 GGAAACCTCA CTCCTTTTGT CACCATAACA TGCATCGTTG CTGCCATGGG
    VfSTP1 GGAAACTTAA CTCCTTTCGT CACCATAACA TGTGTTGTTG CAGCCATGGG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP GGAAACCTTA CTCTTTATGT TACTGTAACA TGTGTCGTTG CAGCCATGGG
      VvHT GGTAACCTTA CTCCTTACGT GACTGTGACA TGTGTTGTTG CAGCCATGGG
     tMST2 GGCAACCTTA CTCCTTTTGT CACTGTAACA TGTATTGTTG CTGCCATGGG
Full-CpHT1 GGTAATCTTA CCCCATTCGT CACTGTAACA TGTATCGTCG CCGCCATGGG
    AtSTP1 GGCAAACTCA CTCCCTTTGT TCTCTTCACT TGCGTTGTTG CTGCCATGGG
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 GCTAAAATCA CACCAATTGT TATCATATCT TGTATTATGG CTGCTACCGG

          201                                                       250
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 TGGTCTCATT TTCGGTTACG ATATTGGAAT TTCTGGAGGT GTGACATCAA
     MtST1 TGGTTTGATC TTTGGCTACG ATATTGGAAT TTCAGGTGGT GTGACGTCCA
    VfSTP1 TGGTTTGATC TTTGGTTACG ATATAGGAAT TTCAGGTGGT GTTACTTCAA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TGGCTTGATC TTTGGTTACG ATATTGGGAT TTCTGGTGGA GTTACGTCCA
      VvHT TGGTTTGATC TTTGGTTACG ATATTGGAAT TTCTGGTGGG GTCACGTCCA
     tMST2 TGGTTTGATC TTTGGTTACG ATATTGGGAT TTCTGGTGGA GTTACGTCCA
Full-CpHT1 TGGACTGATC TTTGGATACG ATATTGGGAT CTCAGGGGGA GTGACGTCAA
    AtSTP1 CGGTCTCATC TTCGGATACG ATATCGGAAT CTCCGGTGGT GTGACGTCTA
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 AGGTCTCATG TTTGGTTATG ATGTTGGAGT TTCTGGTGGT GTTACATCAA
```

Figure 5.17. (continue).

```
      251                                                        300
       LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      NtMST1 TGGACTCATT CTTGAGCAGA TTTTTCCCAT CTGTGTTCAG GAAGCAAAAG
       MtST1 TGGATCCGTT TCTGAAGAAA TTTTTTCCGG CGGTGTACCG GAAAAAGAAC
      VfSTP1 TGAATCCGTT TCTTGAGAAA TTTTTTCCGG CGGTGTACCG GAAGAAAAAC
      PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       RcSCP TGGATTCATT CTTGAAGAAG TTCTTTCCTT CAGTTTACCG GAAGAAGAAA
        VvHT TGGCTCCGTT CTTGCAGAAG TTCTTCCCTT CTGTGTACCG GAAGGAGGCT
       tMST2 TGCCATCATT CTTGAAGAAG TTTTTTCCAT CTGTTTACCG CAAACAGCAA
  Full-CpHT1 TGAACTCGTT TCTGAAGGAA TTTTTCCCGG CGGTTTTCCG GAAAAAGGAA
      AtSTP1 TGCCGTCTTT CCTCAAGCGA TTCTTCCCGT CGGTGTATCG GAAACAACAA
       LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       LeHT2 TGGATCCATT TTTAAAAAAA TTCTTTCCAA CAGTTTATAA GAGAACAAAG

             301                                                  350
       LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      NtMST1 GCAGATGATT CAACAAATCA ATACTGCAAA TTTGACAGCC AAACATTGAC
       MtST1 AAGGACAAAT CGACAAACCA GTACTGTCAA TATGACAGTC AAACATTGAC
      VfSTP1 GCGCAACATT CGAAGAATCA GTACTGTCAA TACGACAGTG AGACACTGAC
      PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       RcSCP GCGGATGAAT CGTCAAACCA GTACTGTCAA TATGATAGTC AGACACTGAC
        VvHT TTGGACAAGT CCACGAATCA GTACTGTAAG TTTGATAGTG AGACACTAAC
       tMST2 GAAGACGCCA CATCAAACCA GTACTGCCAA TATGACAGTC AAACACTAAC
  Full-CpHT1 GAGGTATCGT CGACTAACCA GTACTGTCAG TACGACAGTC CGACACTTAC
      AtSTP1 GAGGACGCGT CAACGAACCA GTACTGTCAG TACGATAGCC CGACGCTAAC
       LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       LeHT2 GAGCCAGGAT TAGACAGTAA TTACTGTAAA TACGATAATC AAGGGTTACA

             351                                                  400
       LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      NtMST1 GATGTTCACG TCGTCATTGT ACTTGGCTGC TCTTTTGTCG TCTCTGGTGG
       MtST1 GATGTTTACA TCGTCGTTGT ATCTGGCTGC CCTTTTGTCC TCGTTGGTAG
      VfSTP1 CTTGTTCACA TCCTCGTTGT ACCTGGCCGC GCTTTTGTCG TCGGTGGTTG
      PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       RcSCP TATGTTCACA TCTTCACTGT ACTTGGCTGC TTTAATTGCT TCTCTTGTAG
        VvHT GTTGTTCACA TCGTCGCTTT ATCTGGCTGC TCTTCTCTCG TCGCTGGTGG
       tMST2 CATGTTCACT TCTTCTCTGT ACTTGGCTGC TTTATTGGCA TCGCTTGTGG
  Full-CpHT1 GTTGTTTACA TCATCGCTGT ATCTGGCGGC GCTTGTGGCG TCGCTGGTTG
      AtSTP1 GATGTTCACA TCGTCTCTAT ATCTAGCGGC GCTAATTTCG TCGCTGGTGG
       LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
       LeHT2 ATTATTTACT TCATCGTTAT ACCTAGCTGG TTTAACGGCG ACGTTTTTTG
```

Figure 5.17. (continue).

```
          401                                                        450
      LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     NtMST1 CATCTACTGT CACCAGAAAA CTTGGACGGA GACTTTCTAT GCTCTGTGGA
      MtST1 CTTCCACCAT AACTCGTAGG TTTGGTCGGA AACTTTCCAT GCTTTTCGGA
     VfSTP1 CTTCAACGAT CACTAGAAGG TTTGGTCGGA AACTCTCCAT GCTTTTCGGA
     PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      RcSCP CATCTACTAT TACGAGAAAA TTTGGTAGGA AACTCTCTAT GCTTTTTGGC
       VvHT CCGCGACGGT GACCCGAAAG TTCGGGAGAA AGCTGTCAAT GCTATTCGGA
      tMST2 CATCCATAGT TACTCGTAAA TTTGGAAGGA AACTCTCCAT GTTGTTTGGA
 Full-CpHT1 CGGCGACGGT GACAAGAAAG TTCGGTCGGA AACTGTCGAT GCTGTTTGGC
     AtSTP1 CTTCCACCGT GACAAGAAAG TTCGGACGGC GGCTCTCGAT GCTCTTCGGC
      LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      LeHT2 CTTCTTACAC TACAAGAAAA CTTGGCCGGA GATTAACTAT GTTAATCGCC

          451                                                        500
      LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     NtMST1 GGTGTCCTCT TCTGTGCTGG AGCTTTGATC AATGGCTTTG CTCAGAATGT
      MtST1 GGATTACTTT TCCTTGTCGG TGCTCTTATT AATGGCTTTG CTAATCATGT
     VfSTP1 GGCTTGCTTT TTCTGGTCGG TGCTCTCATT AATGGCCTTG CTCAAAACGT
     PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      RcSCP GGTGTACTCT TTTGTGCTGG AGCTATCATC AATGGTGCGG CTAAAGCAGT
       VvHT GGACTGCTCT TTTGTGCCGG TGCCATCATC AATGGCGCTG CTAAAGCTGT
      tMST2 GGTGTTCTCT TTTGTGCTGG TGCCATCATC AATGGTTTTG CTCAAGCTGT
 Full-CpHT1 GGCGTCCTGT TCTGCGCCGG TGCCATCATT AATGGCTTCG CTAAAGCTGT
     AtSTP1 GGCATACTCT TCTGCGCCGG AGCTCTCATC AATGGTTTCG CCAAACATGT
      LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      LeHT2 GGTTGTTTTT TCATTATTGG AGTTGTACTA AATGCTGCTG CTCAAGATTT

          501                                                        550
      LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     NtMST1 TGCTATGCTC ATTGTTGGTC GTATTTTACT AGGTTTTGGT ATTGGATTCG
      MtST1 TTGGATGTTG ATCGTGGGTC GGATCTTGCT CGGGTTTGGT ATCGGGTTTG
     VfSTP1 TGCGATGTTG ATCGTCGGTC GGATCTTGCT CGGATTCGGT ATCGGGTTTG
     PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      RcSCP CTGGATGTTG ATTCTTGGTA GAATTTTGCT TGGTTTTGGT ATTGGGTTTG
       VvHT TTGGATGTTG ATTGTCGGTC GTATACTGCT GGGTTTTGGT ATTGGGTTTG
      tMST2 ATGGATGTTG ATTCTCGGTA GAATCTTGCT TGGTTTTGGT ATAGGTTTTG
 Full-CpHT1 TTGGATGTTG ATTCTCGGCA GAATTTTGTT GGGTTTTGGC ATCGGTTTTG
     AtSTP1 TTGGATGCTC ATCGTCGGTC GTATCTTGCT TGGTTTCGGT ATCGGTTTCG
      LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
      LeHT2 AGCTATGCTT ATTATTGGAA GAATTCTCCT TGGTTGTGGC GTTGGTTTCG
```

Figure 5.17. (continue).

```
          551                                                       600
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 CCAATCAGTC TGTTCCACTA TACCTATCTG AAATGGCTCC ATACAAGTAC
     MtST1 CTAATCAGCC TGTGCCATTG TACCTCTCTG AGATGGCTCC TTACAAGTAT
    VfSTP1 CGAATCAGTC TGTGCCATTA TACTTGTCTG AGATGGCTCC ATACAAGTAC
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP CCAATCAGTC TGTGCCGCTC TACCTCTCTG AGATGGCTCC ATACAAATAC
      VvHT CCAATCAGTC TGTGCCGCTC TACCTCTCTG AGATGGCTCC ATACAAATAC
     tMST2 CCAATCAGTC TGTGCCACTC TACCTCTCTG AGATGGCTCC ATACAAGTTC
Full-CpHT1 CCAATCAGTC TGTACCACTC TACCTCTCTG AGATGGCTCC TTACAGATAT
    AtSTP1 CTAATCAGGC TGTGCCACTG TACCTCTCTG AGATGGCTCC ATACAAATAC
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 CCAATCAGGC TGTTCCATTA TTTTTATCAG AAATAGCACC TACAAGAATT

          601                                                       650
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 AGAGGAGCAC TCAACCTAGG TTTTCAACTG TCCATTACAA TTGGTATACT
     MtST1 AGAGGAGCAT TGAATATTGG GTTTCAATTA TCAATTACAA TTGGTATACT
    VfSTP1 AGAGGAGCGT TGAATATTGG ATTTCAATTG TCAATTACAA TTGGAATACT
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP AGAGGAGCAC TGAACATTGG TTTCCAGTTA TCAATTACAA TTGGTATCCT
      VvHT AGAGGAGCCC TCAACATTGG CTTCCAATTA TCCATCACAA TTGGTATTCT
     tMST2 AGAGGTGCAC TCAACATTGG TTTTCAATTA TCAATCACAA TTGGTATCCT
Full-CpHT1 AGAGGAGCAT TAAACATTGG ATTCCAATTG TCCATCACAA TTGGTATTCT
    AtSTP1 AGAGGAGCTT TAAACATTGG TTTCCAGCTC TCAATTACAA TCGGAATCCT
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 CGTGGAGGAC TTAACATTTT GTTCCAACTT AACGTAACTA TTGGTATTCT

          651                                                       700
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 TGTAGCAAAT GTGTTGAACT ATTTCTTTGC CAAGATTCA. .....TTGGG
     MtST1 TGTGGCCAAT GTGTTGAATT ACTTTTTTGC CAAAATCAAA GGTGGATGGG
    VfSTP1 TGTGGCCAAT ATTTTGAACT ACTTTTTTGC CAAAATCAAA GGTGGATGGG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TGTAGCCAAT GTATTGAATT ACTTCTTTGC CAAGATTAAG GGTGGTTGGG
      VvHT TGTGGCCAAT ATATTGAACT ACTTCTTTGC AAAGATCAAG GGGGGTTGGG
     tMST2 TGTCGCCAAT GTGTTAAACT ATTTCTTTGC TAAAATCCAT GGTGGTTGGG
Full-CpHT1 TGTTGCCAAT GTATTGAATT TCTTCTTTGC AAAAATCAAA GGAGGTTGGG
    AtSTP1 CGTCGCCGAA GTGCTAAACT ACTTCTTCGC CAAGATCAAA GGCGGTTGGG
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 TTTTGCCAAT CTCGTCAACT ACGGAACAGC CAAGATTAGT GGAGGATGGG
```

Figure 5.17. (continue).

```
          701                                                       750
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 GATGGAGATT AAGCTTAGGA GGTGCTATGG TACCTGCATT GATCATCACA
     MtST1 GATGGAGATT GAGTTTAGGT GGTGCTATGG TCCCAGCACT TATAATAACA
    VfSTP1 GATGGAGATT GAGTTTAGGT GGTGCTATGG TCCCTGCACT TATAATAACC
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP GTTGGAGGCT GAGTCTTGGT GGTGCTATGG TCCCTGCCCT CATCATTACA
      VvHT GATGGAGATT GAGCTTGGGT GGCGCTGTGG TCCCTGCGCT CATCATCACC
     tMST2 GATGGAGATT GAGTTTGGGT GGTGCTATGG TTCCTGCCCT TATAATCACA
Full-CpHT1 GATGGAGACT GAGCTTAGGA GGTGCAGTAG TTCCAGCTCT AATCATCGCC
    AtSTP1 GATGGCGGCT CAGTCTCGGA GGCGCGGTGG TTCCTGCCTT GATCATAACC
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 GATGGAGATT ATCATTAGGG TTAGCTGGAT TTCCAGCAGT ATTGTTGACT

          751                                                       800
     LeHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    NtMST1 ATAGGCTCAC TTTTCCTTCC CGAGACACCA AACTCCATGA TCGAACGTGG
     MtST1 ATTGGATCAT TAGTCCTTCC CGACACCCCT AACTCAATGA TCGAACGTGG
    VfSTP1 ATTGGTTCGC TAATCCTACC CGACACGCCA AATTCCATGA TCGAGCGTGG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP GTTGGATCAT TAGTCCTTCC AGATACACCA AACTCCATGA TTGAACGGGG
      VvHT GTCGGGTCCC TTGTCCTCCC GGACACACCC AACTCCATGA TCGAGCGTGG
     tMST2 GTTGGATCAC TAGTCCTTCC TGATACTCCA AACTCCATGA TCGAACGTGG
Full-CpHT1 ATTGGATCGT TAATCCTCCC CGATACACCC AACTCCATGA TCGAACGAGG
    AtSTP1 ATCGGCTCCC TCGTCCTCCC TGACACTCCC AATTCAATGA TCGAGCGTGG
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 TTGGGCGCAT TATTTGTTGT TGAAACCCCA AACAGTTTGA TTGAAAGAGG

          801                                                       850
     LeHT1 ~AACCACGAC GAAGCCAAAG CTCGATTGAA GAGAATTAGG GGAATTGAAG
    NtMST1 CAATCACGAC GAAGCCAAAG CTAGGCTTAA AAGAATCAGA GGCATTGATG
     MtST1 TGATCGCGAT GGAGCTAAAG CTCAACTTAA GAGAATTCGC GGCATTGAAG
    VfSTP1 GGATCGAGAT GGTGCCAAGG CACAGCTTAA GAGAATTCGC GGAGTTGAAG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP CCAGCACGAA GAAGCCAGAG CACATCTAAA GAGAGTTCGC GGTGTAGAAG
      VvHT CCAGCACGAG GGAGCGAAAA CAAAACTGAG AAGAATCCGG GGTGTCGATG
     tMST2 CCAGCATGAT GAGGCTAGAG AAAAATTGAG AAGAGTTCGT GGTGTTGATG
Full-CpHT1 CCAGGTAGAA GCAGCTAAAG AGAAATTAAG GAGAATTCGG GGTGTCAACA
    AtSTP1 CCAACACGAA GAAGCCAAAA CCAAGCTCAG ACGAATCCGT GGTGTCGATG
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 TTACTTAGAA GAAGGCAAAG AAGTACTTCG AAAAATCAGA GGTACCGACA
```

Figure 5.17. (continue).

```
          851                                                       900
     LeHT1 ATGTAGATGA AGAGTTCAAT GATTTGGTTA TTGCTAGTGA AGCTTCTAGG
    NtMST1 ATGTAGACGA AGAGTTCAAT GATTTAGTCG TGGCGAGTGA GGCTTCTAGG
     MtST1 ATGTTGATGA AGAGTTTAAT GACCTCGTAG CAGCTAGTGA GGCCTCAATG
    VfSTP1 ATGTTGATGA AGAGTTTAAT GATCTTGTGG CTGCTAGTGA AACGTCGATG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP ATGTTGATGA GGAGTTTACT GACCTTGTTC ATGCTAGTGA AGATTCAAAG
      VvHT ATGTTGAAGA GGAATTCAAT GACCTTGTTG TAGCCAGTGA GGCCTCCAAG
     tMST2 ATGTTGATGA GGAGTTTAAC GACCTTGTTG CTGCTAGTGA AGCTTCAATG
Full-CpHT1 ACGTGGACGA AGAGTTAAAA GACTTAGTTG CAGCAAGTGA AGCTTCGAAA
    AtSTP1 ACGTCAGCCA AGAGTTTGAC GATTTGGTCG CCGCTAGTAA AGAGTCGCAG
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 ACATTGAACC TGAATTCTTG GAACTTGTTG AGGCTAGTCG TGTTGCTAAA

          901                                                       950
     LeHT1 AAAATTGAAC ATCCCTGGAG GAACTTGTTG CAAAAGAAAT ATAGACCACA
    NtMST1 AAAATTGAGA ACCCTTGGAG AAATTTGTTG CAAAGGAAAT ATAGGCCACA
     MtST1 CAAGTTGAAA ACCCTTGGAG AAATTTGTTG CAGAGGAAAT ATAGACCTCA
    VfSTP1 CAGGTTGAAA ATCCTTGGAG GAATTTGTTG CAGAGAAAAT ATAGACCTCA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP AAAGTTGAAC ATCCTTGGAG GAATTTGTTA CAGAGGAAAT ACAGGCCTCA
      VvHT CTTGTTGAGC ACCCCTGGAG AAATCTCTTG CAGAGGAAGT ACAGGCCACA
     tMST2 AAAGTAGAAC ATCCATGGAG AAATTTATTG CAAAGAAAGT ACAGGCCTCA
Full-CpHT1 TTGGTAGAAC ATCCATGGAG AAACTTGTTA CAAAGAAAAT ACAGGCCTCA
    AtSTP1 TCGATAGAGC ACCCGTGGAG AAACCTCCTC CGCCGCAAGT ACCGACCACA
     LeHT3 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ AAAAGGAGAA ACAGGCCTCA
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 CAAGTCAAAC ACCCTTTCAG AAATTTACTC CAACGTAAAA ATAGACCTCA

          951                                                      1000
     LeHT1 TCTTACAATG GCAATTATGA TCCCATTTTT CCAACAACTT ACTGGAATCA
    NtMST1 TCTCACAATG GCAATTATGA TCCCATTTTT CCAGCAACTT ACTGGAATCA
     MtST1 GCTTACTATG GCTGTATTGA TACCATTCTT CCAACAATTT ACAGGCATCA
    VfSTP1 GCTTACTATG GCTGTGTTGA TTCCGTTCTT CCAACAGTTT ACTGGAATTA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TCTCTCAATG GCCATTGCAA TTCCGTTCTT TCAGCAACTC ACCGGCATTA
      VvHT CCTCACAATG GCCATCCTCA TTCCCTTCTT CCAGCAGCTT ACCGGGATTA
     tMST2 CATAACCATG GCAGTGATGA TTCCATTCTT TCAGCAACTT ACTGGCATTA
Full-CpHT1 TCTCACCATG GCTATCATGA TCCCATTCTT CCAGCAGCTA ACTGGAATTA
    AtSTP1 TCTCACAATG GCCGTTATGA TTCCGTTCTT TCAACAGCTA ACCGGAATCA
     LeHT3 GTTAATCATG GCGATAATGA TGCCGACTTT TCAGATACTT ACTGGCATTA
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 ATTAATCATC TCTGTTGCCC TCCAGATATT CCAACAATTT ACAGGAATCA
```

Figure 5.17. (continue).

```
          1001                                                      1050
     LeHT1 ACGTGATTAT GTTTTATGCA CCTGTGTTGT TTAAAACCAT TGGTTTTGGT
    NtMST1 ATGTGATTAT GTTCTATGCA CCAGTTTTGT TTAAGACTAT TGGTTTTGGT
     MtST1 ATGTTATCAT GTTTTATGCA CCTGTGCTAT TTAATTCCAT TGGGTTTAAG
    VfSTP1 ACGTTATTAT GTTTTACGCG CCTGTGTTGT TTAACTCGAT TGGGTTTAAG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP ATGTGATCAT GTTCTATGCT CCTGTTTTGT TCGATACTAT TGGATTCGGT
      VvHT ATGTCATTAT GTTTTATGCC CCTGTTCTCT TCAAAACTAT TGGCTTTGCG
     tMST2 ATGTCATTAT GTTTTACGCT CCTGTTTTGT TCAACACAAT TGGTTTTGGT
Full-CpHT1 ATGTCATCAT GTTTTATGCT CCTGTTTTAT TCAACACAAT TGGGTTTGGC
    AtSTP1 ATGTGATTAT GTTTTACGCT CCGGTTTTGT TCAACACCAT TGGTTTCACG
     LeHT3 ACATCATACT TTTTTATGCC CCAGTATTGT TTCAGAGTAT GGGGTTTAAA
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 ACGCTATTAT GTTCTACGCA CCTGTTTTAT TTTCAACACT AGGGTTTGGA

          1051                                                      1100
     LeHT1 ACTGATGCTT CACTTATGTC TGCTGTGATC ACTGGTGGAA TCAATGTCAT
    NtMST1 GCTGATGCTT CCCTTATGTC TGCTGTTATT ACTGGTGGAG TCAATGTACT
     MtST1 GACGATGCTT CACTTATGTC GGCTGTCATC ACCGGTGTTG TTAATGTTGT
    VfSTP1 GATGATGCTT CACTTATGTC AGCTGTTATC ACCGGTGTTG TTAACGTTGT
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP AGTGATGCTG CACTCATGTC TGCTGTGATC ACTGGTCTTG TTAATGTTTT
      VvHT GATGATGCTT CCCTGATGTC TGCTGTGATA ACCGGGCGGG TTAATGTTCT
     tMST2 AGCAATGCTT CGCTCATGTC TGCTGTGATC ACTGGTGTTG TTAATGTCGT
Full-CpHT1 AGTGACGCCT CTCTCATGTC TGCTGTAATT ACCGGAATTG TAAATGTCGG
    AtSTP1 ACCGATGCTT CTCTCATGTC CGCTGTGGTC ACTGGCTCGG TTAACGTTGG
     LeHT3 AGAGCAGCCT CTCTGTATTC CTCTGCTTTG ACTGGTGCAG TTCTTGCTTC
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 AACAGCGCAG CTCTTTACTC AGCTGTCATC ACCGGAGCAG TCAACGTTCT

          1101                                                      1150
     LeHT1 TGCCACTATT GTTTCTATTT ACTATGTTGA TAAATTAGGA AGAAGATTCT
    NtMST1 TGCAACTGTT GTTTCTATTT ACTATGTTGA TAAATTGGGA AGAAGATTCT
     MtST1 TGCTACTTGT GTCTCAATTT ATGGAGTTGA TAAGTGGGGT AGGAGAGCCC
    VfSTP1 TGCTACTTGT GTCTCAATTT ATGGAGTTGA TAAGTGGGGG AGAAGAGCTC
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TGCAACAATG GTCTCAATTT ATGGTGTTGA TAAGTGGGGA AGGAGGTTCC
      VvHT TGCAACCATA GTTTCAATCT ACGGTGTTGA TAAGTGGGTA AGAAGGTTTC
     tMST2 TGCTACCATG GTTTCAATCT ATGGCGTCGA CAAGTGGGGG AGAAGGTTTC
Full-CpHT1 TGCAACTTTG GTTTCAATCT ATGGAGTCGA TAAATGGGGA AGACGATTCC
    AtSTP1 CGCTACGCTT GTTTCTATCT ACGGTGTTGA CAGATGGGGA CGTCGGTTTC
     LeHT3 ATCTACACTT TTATCAATGG CCACTGTCGA TAGATGGGGT CGAAGAGTTC
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 CTCCACCGTT GTCTCAGTCT ACTCCGTTGA CAAGCTCGGA CGACGAGTCC
```

Figure 5.17. (continue).

```
          1151                                                      1200
     LeHT1 TGTTTCTTGA AGGTGGAATT CAAATGCTCT TTTCCCAAAT AGCCGTGGCA
    NtMST1 TGTTCCTTGA AGGTGGCATT CAAATGCTCA TCTGCCAAAT AGCGGTGTCA
     MtST1 TTTTCCTTGA AGGTGGTGCT CAAATGCTCA TATGCCAGGT TGCAGTAGCA
    VfSTP1 TTTTTCTCGA AGGTGGTGTT CAAATGCTTA TTTGTCAGGT TGCAGTTGCA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TTTTCCTTGA GGGTGGAGTT CAAATGTTGA TTTGCCAGGC AATTGTTGCA
      VvHT TTTTCCTTGA GGGTGGCACT CAAATGCTCA TATGTCAGGT TATTGTGGCA
     tMST2 TTTTCCTTGA GGGTGGTTTT CAGATGTTGA TATGCCAGGC AGTCGTAGCA
Full-CpHT1 TTTTCCTCGA GGGAGGAGTT CAAATGTTAA TATGCCAGAT TGTGGTAGCA
    AtSTP1 TCTTTCTTGA AGGTGGTACA CAAATGCTTA TATGCCAGGC TGTGGTTGCA
     LeHT3 TTCTTATTAC CGGTGGAATC CAAATGATCA TCTGTCAGGT TATTGTTGCG
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 TCCTCCTGGA AGCTGGGGTC CAAATGTTAC TATCTCAAAT AATAATCGCT

          1201                                                      1250
     LeHT1 ATTTTGATAG CAATAAAGTT TGGAGTAAAT GGAACTCCAG GGGAATTACC
    NtMST1 ATTTGCATAG CTATAAAATT TGGAGTGAAT GGAACTCCAG GGGATTTACC
     MtST1 GCTGCAATTG GGGCCAAATT TGGAACAAGT GGAAACCCTG GTAATTTACC
    VfSTP1 GTTTCAATTG CGGCCAAGTT TGGAACAAGT GGAGAACCTG GTGATTTACC
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP GCCTGCATTG GTGCTAAGTT TGGAGTAGAT GGAGCTCCCG GTGACTTGCC
      VvHT ACGTGCATTT TGGTTAAATT CGGAGTGGAT GGAGAACCTT GGTGCTTGCC
     tMST2 GCTTGTATTG GTGCCAAGTT TGGAGTTAAC GGGAACCCTG GTGAATTGCC
Full-CpHT1 GCCTCCATTG GAGCTAAATT TGGGATCAAT GGCAACCCTG GAGATTTACC
    AtSTP1 GCTTGCATAG GGGCCAAGTT TGGGGTAGAC GGGACCCCTG GTGAGCTACC
     LeHT3 ATAATCTTGG GACTCAAATT .........T GGAAGTGACA AGGAGCTATC
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 ATAATCCTAG GTATCAAAGT T......ACT GATCATTCAG ACAACCTTAG

          1251                                                      1300
     LeHT1 AAAATGGTAT GCAATAGTGG TTGTGATATT CATTTGTGTA TATGTTGCTG
    NtMST1 AAAGTGGTAC GCGATAGTAG TGGTGATATT CATCTGTGTT TATGTAGCTG
     MtST1 AGAATGGTAT GCTATAGTAG TTGTGCTCTT CATTTGCATT TACGTAGCAG
    VfSTP1 AAAGTGGTAT GCTATAGTAG TTGTGCTTTT CATATGCATT TACGTTGCTG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP ACAATGGTAT GCAGTCGTTG TGGTGCTTTT CATTTGCATT TACGTATCTG
      VvHT CAAGTGGTAT GCCATAGTTG TGGTGCTGTT CATTTGCGTC TATGTTTCAG
     tMST2 CAAGTGGTAT GCTATTGTTG TGGTGCTTTT CATTTGCATT TACGTTGCGG
Full-CpHT1 AAAATGGTAT GCAATTGTAG TGGTGCTATT CATCTGTATT TACGTGGCCG
    AtSTP1 AAAGTGGTAT GCTATAGTGG TTGTAACGTT CATTTGCATC TATGTGGCGG
     LeHT3 AAGAGGTTAC TCGATTATAG TAGTTGTTTT CATTTGCCTC TTTGTAGCGG
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 CCATGGTTGG GGAATCTTCG TAGTAGTTCT GATCTGTACA TATGTATCGG
```

Figure 5.17. (continue).

```
          1301                                                    1350
     LeHT1 GATTCGCTTG GTCATGGGGT CCTCTTGGAT GGCTCGTACC TAGTGAAATT
    NtMST1 GATTTGCTTG GTCCTGGGGA CCTCTAGGAT GGTTGGTACC TAGTGAAATT
     MtST1 GATTTGCTTG GTCATGGGGT CCTCTTGGTT GGTTGGTTCC TAGTGAGATT
    VfSTP1 GATTTGCTTG GTCATGGGGT CCTCTTGGTT GGTTGGTGCC TAGTGAGATT
    PpSTP1 ~~~~~~~~~~ ~~~~TGGGGG CCGCTGGGGT GGCTCGTCCC CAGTGAAATC
     RcSCP GATTCGCCTG GTCTTGGGGT CCCCTGGGAT GGCTGGTGCC AAGTGAAATC
      VvHT GGTTTGCATG GTCCTGGGGA CCTCTAGGTT GGTTGGTCCC TAGTGAAATT
     tMST2 GATTTGCTTG GTCTTGGGGC CCTCTTGGTT GGTTGGTGCC AAGTGAGTTT
Full-CpHT1 GATTTGCGTG GTCTTGGGGG CCTCTCGGGT GGCTCGTGCC GAGTGAAATC
    AtSTP1 GTTTTGCGTG GTCGTGGGGC CCACTAGGGT GGTTAGTACC GAGTGAAATC
     LeHT3 CGTTTGGATA CTCATGGGGG CCTCTTGGAT GGACCGTGCC AAGTGAAATT
    PpSTP2 ~~~~~~~~~~ ~~~~TGGGGG CCGCTGGGGT GGACAGTGCC AAGCGAGATA
     LeHT2 CTTTCGCATG GTCATGGGGC CCACTTGGAT GGTTAATTCC TAGCGAGACG

          1351                                                    1400
     LeHT1 TTCCCACTGG AAATTCGATC AGCTGCACAA AGTATCAATG TCTCAGTGAA
    NtMST1 TTCCCACTTG AAATTCGATC AGCTGCTCAA AGTATCAATG TTTCAGTGAA
     MtST1 TTCCCATTGG AGATTCGTTC TGCAGCTCAA AGTGTAAACG TATCTGTGAA
    VfSTP1 TTTCCATTGG AGATCCGTTC TGCTGCGCAG AGTGTCAACG TGTCGGTCAA
    PpSTP1 TTCCCACTGG AAATCCGATC AGCTGCACAG AGCGTCAACG TTTCAGTGAA
     RcSCP TTCCCACTGG AAATTCGGTC TGCTGCACAA AGTGTGAATG TTTCTGTCAA
      VvHT TTCCCCCTGG AAATCCGATC TGCTGCACAG AGTGTAAACG TCTCCGTTAA
     tMST2 TTCCCACTGG AAATCCGATC AGCTGCACAA AGTATCAGTG TGTCTGTCAA
Full-CpHT1 TTCCCTCTTG AAATCAGATC AGCAGCTCAG AGTATCAATG TGTCGGTGAA
    AtSTP1 TTCCCGTTGG AGATAAGGTC GGCGGCGCAG AGTATCACCG TGTCCGTGAA
     LeHT3 TTCCCTTTAG AGACGAGATC AGCAGGCCAA AGTATCACAG TTACTGTGAA
    PpSTP2 TTCCCACTGG AAACCCGATC AGCCGGACAA AGCATTACAG TGGCTGTGAA
     LeHT2 TTCCCTTTGG AAACTCGTTC AGCTGGTCAA AGTGTAACAG TGTGTGTTAA

          1401                                                    1450
     LeHT1 CATGATCTTC ACATTTGCAG TAGCACAAGT TTTCTTAACA ATGTTGTGTC
    NtMST1 CATGATCTTC ACATTTATAG TGGCACAAGT ATTCTTGACA ATGTTGTGTC
     MtST1 CATGCTTTTC ACCTTCTTAG TTGCACAAGT TTTCTTGATA ATGCTTTGTC
    VfSTP1 CATGCTCTTC ACCTTCTTAG TTGCACAGAT TTTCTTGACC ATGCTTTGTC
    PpSTP1 CATGATCTTC ACTTTCCTCG TGGCTCAAAT CTTCTTGACG ATGCTCTGCC
     RcSCP CATGTTCTTT ACATTTGTAG TAGCTCAAGT ATTCCTGATA ATGCTTTGTC
      VvHT CATGTTTTTC ACATTCATCA TAGCCCAAAT CTTCTTAAAT ATGCTGTGTC
     tMST2 CATGCTTTTC ACTTTCATAG TTGCCCAGAT TTTTCTCACG ATGCTTTGCC
Full-CpHT1 TATGTTCTTC ACATTTATAG TGGCACAAAT ATTTTTGACA ATGCTTTGTC
    AtSTP1 CATGATCTTC ACGTTCATTA TCGCGCAAAT CTTCTTGACG ATGCTTTGTC
     LeHT3 TTTGTTCTTC ACATTTGCGA TAGCACAGTC TTTCCTCTCA CTTTTATGTG
    PpSTP2 CCTTCTGTTC ACCTTCATCA TAGCACAGTC TTTTCTTGCC CTCCTATGTG
     LeHT2 CTTGCTTTTC ACGTTTGTTA TGGCACAAGC ATTTCTCTCA ATGCTTTGTC
```

Figure 5.17. (continue).

```
          1451                                                      1500
     LeHT1 ATTTGAAGTT TGGATTGTTT CTGTTTTTCG CCTTCTTTGT GGTGATTATG
    NtMST1 ATTTGAAGTT TGGATTGTTC CTCTTCTTTG CATTCTTTGT TGTGATTATG
     MtST1 ACATGAAGTT TGGTTTGTTC CTCTTCTTTG CCTTCTTCGT TTTGGTGATG
    VfSTP1 ACATGAAGTT TGGATTGTTC CTCTTCTTTG CCTTCTTTGT GGTGGTGATG
    PpSTP1 ATTTGAAATT TGGGCTATTC CTCTTCTTTG CGTTCTTTGT GTTCGTGATG
     RcSCP ATTTGAAGTT TGGGCTATTC ATCTTCTTTT CATTCTTTGT GTTGATAATG
      VvHT ACATGAAGTT TGGTTTGTTC CTCTCCTTTG CCTTCTTTGT GGTGGTGATG
     tMST2 ACTTGAAGTT TGGGCTATTC CTCTTCTTTG CCTTCTTTGT GGTGCTGATG
Full-CpHT1 ATTTGAAGTT TGGACTGTTC ATTTTCTTTG CCTTCTTTGT GGTTATCATG
    AtSTP1 ATTTGAAGTT TGGGTTATTC CTTGTTTTCG CCTTTTTCGT GGTGGTGATG
     LeHT3 CTATGAGGTT CGGGATTTTC CTGTTTTTCT CCTGTTGGAT TGCTGTCATG
    PpSTP2 CGTTAAAATT CGGGATCTTT CTCTTCTTTG CTGGGTGGAT TACTGTGATG
     LeHT2 ATTTCAAGTA TGGGATATTC TTGTTCTTCT CGGGGTGGAT TTTTGTGATG

          1501                                                      1550
     LeHT1 ACTGTGTTCA TATACTTCTT CTTGCCTGAG ACGAAAAATA TTCCGATAGA
    NtMST1 ACTGTCTTCA TTTACTTCTT CTTGCCTGAG ACAAAGAATA TCCCAATTGA
     MtST1 TCAATCTATG TATTCTTCTT ATTGCCTGAA ACTAAAGGAA TACCAATTGA
    VfSTP1 ACAATTTATA TATACACTAT GCTGCCTGAA ACTAAGGGAA TACCAATTGA
    PpSTP1 TCCATTTTCA TTTACTACTT CTTGCCTGGG ACCAAAGGGG TA~~~~~~~~
     RcSCP TCCATCTTCG TGTACTACTT CTTGCCTGAG ACAAAAGGCA TCCCAATTGA
      VvHT TCCTTCTTCA TTTACTTCTT CTTGCCTGAG ACCAAAGGCA TCCCAATTGA
     tMST2 TCCATCTTTG TCTACTACTT CTTGCCTGAG ACAAAGGGCA TTCCAATCGA
Full-CpHT1 TCAATCTTCA TCTACTATTT CTTGCCGGAG ACAAAGGGAA TCCCCATTGA
    AtSTP1 TCGATCTTTG TATACATTTT CTTGCCGGAG ACGAAAGGGA TTCCGATAGA
     LeHT3 ACGATATTCA TCTATCTGTT CTTGCCTGAA ACGAAGGGAG TTCCGATTGA
    PpSTP2 ACGGTGTTCG TTTACGTGTT CCTGCCTGGA ACCAAGGGGG TA~~~~~~~~
     LeHT2 TCGTTGTTCG TTTTCTTCTT GTTGCCTGAG ACGAAGAATG TTCCTATTGA

          1551                                                      1600
     LeHT1 AGAGATGGTG ATTGTGTGGA AAGAACATTG GTTCTGGTCT AAGTTCATGA
    NtMST1 AGAGATGGTG ATTGTGTGGA AAGAGCATTG GTTCTGGTCT AAGTTCATGA
     MtST1 AGAGATGGAC AGAGTTTGGA AATCACATCC CTTCTGGTCT AGATTTGTTG
    VfSTP1 AGAGATGGAT AGAGTATGGA AATCACACCC ATATTGGTCT AGATTTGTTG
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP AGAGATGGGC CAAGTATGGA AGCAACACTG GTACTGGTCA AGATATGTTG
      VvHT AGAGATGGCT GAAGTATGGA AAAGTCACTG GTTCTGGTCC CGGTATGTCA
     tMST2 AGAGATGGGA CAAGTATGGA AGACTCACTG GTTCTGGTCA AGGTATGTTA
Full-CpHT1 AGAAATGAGT AAGGTCTGGA AGTCTCACTG GTTCTGGTCC AGGTTTGTGG
    AtSTP1 GGAGATGGGT CAAGTGTGGA GGTCACACTG GTATTGGTCA AGGTTTGTGG
     LeHT3 AGAGATGA.. .TGCGTCTTT GGGAAAAGCA TTGGTTCTGG AAGAAGATCG
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 GGAGATGACG GAGAGGGTGT GGAAGCAACA TTGGCTGTGG AAGAGGTTTA
```

Figure 5.17. (continue).

```
          1601                                                      1650
     LeHT1 CTGAAGTTGA TTATCCTGGA ACTAGGAATG GAACTGCTGT TGAAATGGCT
    NtMST1 CTGAAGTGGA CTATCCTGGA ACTAGGAATG GAACAAGTGT TGAAATGTCA
     MtST1 AACATGGTGA TCATGGCAAT GGTGTTGAGA TGGGAAAGGG AGCTCCTAAA
    VfSTP1 AACATG.... ..ATGACAAT GGTGTTGAGA TGGCCAAGGG AGGTGTTAAA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TGGATGAAGA TTATCCTAAT GGAGGGCTTG AAATGGGCAA GGAAGGCCGA
      VvHT ACGATGGTTC TTACAGCGGC GTCGAACTGG TCAAGGAAAA CTACCCTGTT
     tMST2 CCGATGAAGA TTATCCCAAA GCAGGAGGCT ATGAGATGAC CAAAGGAGGG
Full-CpHT1 AAGATGATGG TTATGGCCAT CATGGAAATC TTGAGATGGG AAAAGGCAAC
    AtSTP1 AGGATGGTGA GTATGGGAAT GCGCTTGAGA TGGGCAAGAA CAGTAACCAA
     LeHT3 TTTCGGAGGA TCAACAAGTT AAAAACACCA ATGGACTCAA CCATGCTTGA
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 TGGTGGATGA AGATGATGTT GATATGATTA AAAAGAATGG ACATGCTAAC

          1651                                                      1700
     LeHT1 AAAGGGGGTG CTGGTTACAA AATTGTATGA CTTTAGTTTG GGTTTTTAAA
    NtMST1 AAAGGGAGTG CTGGTTACAA AATAGTATGA CCTAATTAAA G......AAG
     MtST1 AATGTGTAAT TATTATTATT AGTCTTCATT TTATTTTATT TTCATTATTA
    VfSTP1 AATGTATAAT TTATCAGTCT TGCTTTTATT TTTATTTTTG TTACTAATTA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP ATTCCAAAGA ATGTGTAACT AAGCTTGGTG GATTTATTTT AGCTTATATT
      VvHT AAGAATGTAT GA~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     tMST2 CAGGGTCCGA AGAACGTGTA GCTGGGCTTG CTTGGGTTGG ATTTGGTAGT
Full-CpHT1 CTGGGACCAA AGAATGTGtg aTTCATTTTT CTTTTTTTTT TTTTTTGTTT
    AtSTP1 GCTGGAACGA AGCATGTTTG ATTTATCATT GTTTTTAATG AGAGTTTTAA
     LeHT3 GGTAGGTACA TTGTTAACAA GGTGAAAGGT GAATAGAGAA GGCTATCTTG
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 GGATATGATC CCACTTCTCG GTTGTAAAAT AAAAAATAAA AGAATATTAG

          1701                                                      1750
     LeHT1 TTTTTATTTG TTGTTTGTAT AATGTTGTAG TGGGGATGAT ATTGATAATT
    NtMST1 ATGTTTGGAT TTATTTTAAT TTTATTTGTT GTTGTATAAT GTTTTAGTGG
     MtST1 ATTAGTTTTA TTGGTGAAAC ACTAACTATT GGTGTCAACC TCAAGTATCA
    VfSTP1 GTTTTATTTT TATTTTGCTT GTGAAGCTTG GTATTGAATT GTCCTTAAAA
    PpSTP1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     RcSCP TTGAACCCTT TTCCCTTTTC TTCTTACCTT TTGGGATTAG CAGCAGTACA
      VvHT ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     tMST2 CGGCTCTTTT GATTTTTGGT AGCAGTACAT ACCCTACTGA TTTCAATGGT
Full-CpHT1 TTTTCTGGGT TTTAAAGAAT CGGAAGCAGT AAATAGACTT TTTAGTTTGT
    AtSTP1 GAAAGAAAGA AAAAAGATTT GTAATTTCTA ATGTCGTAAA GGAAAAAGTG
     LeHT3 TTCATGTAGA ATTTTGGTCT TCTATATGTA TATGTAAAAG TGTGAACTTT
    PpSTP2 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
     LeHT2 CGGTGTTTTA CGTTGATTTT ATCATAGTAA AGATGCGTAG AAATAAATCA
```

Figure 5.17. (continue).

```
          1751                                                      1800
    LeHT1 ATTAATTAGA TTTGATTGAA ACTGTTTCTA TTGTTTACTT TTGCATAGAA
   NtMST1 GATGATATTG TTAGATTTGA ACTGTTTCTC TTGTCTCTAT TGCATAGAAA
    MtST1 AATGTAATGA AATTGCACTT CAAAATTACG GGATTATTTT TCTCAAAAAA
   VfSTP1 AAAAAAAAAA AAA------- ---------- ---------- ----------
   PpSTP1 ---------- ---------- ---------- ---------- ----------
    RcSCP TACTAATGAT TTCAATATCA AAAATTATGT GGAAATTTTT ----------
      VvHT ---------- ---------- ---------- ---------- ----------
    tMST2 CAAGAATTAT ATGAAATTGG TCTTAAATTT ATCTCGAATT GAAAGGGTCT
Full-CpHT1 TTGTATTACG GAATTACGAA AAACTCTCGA TAATACAGAT CAATTCTATT
   AtSTP1 TATTAGCCTA GATATTTATT GGTGTTTATA TAATTCAATA CCACATGAAG
    LeHT3 TTACTTTTAA GAATGTCACT TTCTATGAAT ACACGAGTGA AATGATATTT
   PpSTP2 ---------- ---------- ---------- ---------- ----------
    LeHT2 AACATGCAAG AGAGAGATCA GAGTCTACTG TCTAGTCTAG TGATATCGAT

          1801                                                      1850
    LeHT1 AAAAATACAT AACTTTGTTC AATAGAAAAT TTGGCAAAGC AACTGTGTAG
   NtMST1 ATACATACTT TACTGTGTTC AATTTCAACC TCTCAACAAT TAATATAAGA
    MtST1 AAAAAAAAAA AAAAAAAAAA AAAAAAA--- ---------- ----------
   VfSTP1 ---------- ---------- ---------- ---------- ----------
   PpSTP1 ---------- ---------- ---------- ---------- ----------
    RcSCP ---------- ---------- ---------- ---------- ----------
      VvHT ---------- ---------- ---------- ---------- ----------
    tMST2 TGAGCTTTAT GGTCAGCAAA AATCCCAATG TATAGACAAG AATCTATTGT
Full-CpHT1 CAATAATCAT AGTAAGTTAT GTTCTAAAAA AAAAAAAAAA AAAAAAAAAA
   AtSTP1 AAATTATGCA TATGATTCTT CGTTAATTGT CCGTAATTGT TATACTCTTT
    LeHT3 TGATAAAAAA AAAAAAAAAA ---------- ---------- ----------
   PpSTP2 ---------- ---------- ---------- ---------- ----------
    LeHT2 AAGTGTTTCT CACAAAATAG GTGTGCATTT GGATTTTTTT TACTGCTTGT

          1851                                                      1900
    LeHT1 CCTTTTGTTT GTTTTGTTGG ATGTACTATT TAGTAGTTCT AGTCTTTTAG
   NtMST1 TTTGTCAGAG CAAAA----- ---------- ---------- ----------
    MtST1 ---------- ---------- ---------- ---------- ----------
   VfSTP1 ---------- ---------- ---------- ---------- ----------
   PpSTP1 ---------- ---------- ---------- ---------- ----------
    RcSCP ---------- ---------- ---------- ---------- ----------
      VvHT ---------- ---------- ---------- ---------- ----------
    tMST2 GTGGAAACCT TATTTAAACC GATTTCAATG CAAAAATAAT TTCTTCAGGA
Full-CpHT1 AA-------- ---------- ---------- ---------- ----------
   AtSTP1 ACTTAAACCA AGTGTTTTCT CTTTGAAAAA AAAAAAAAAA AAAAAA----
    LeHT3 ---------- ---------- ---------- ---------- ----------
   PpSTP2 ---------- ---------- ---------- ---------- ----------
    LeHT2 TTTCCTCCTT TATATCGTTC AATGTAGTAG AAATTGTTTA GCCAAAGTGT
```

Figure 5.17. (continue).

```
          1901                                                      1950
   LeHT1  TGTAAGATCT TTATTATAAA AAATTATGTT CATAAGCTGT ATAAAAAAAA
  NtMST1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   MtST1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  VfSTP1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PpSTP1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   RcSCP  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
    VvHT  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   tMST2  AAAAAAAAA~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
Full-CpHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  AtSTP1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   LeHT3  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  PpSTP2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
   LeHT2  AAGAGATGCA ATATGGTGAT AATTTATGTA ACCATATAGC CTATGTATAT

          1951                                                1995
   LeHT1  AAAAAAAAAA AA~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
  NtMST1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
   MtST1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
  VfSTP1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
  PpSTP1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
   RcSCP  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
    VvHT  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
   tMST2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
Full-CpHT1 ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
  AtSTP1  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
   LeHT3  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
  PpSTP2  ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~
   LeHT2  TAGAGATTTG GTATAATTCA TTTTGGGAAA AAAAAAAAAA AAAAA
```

Figure 5.17. (continue).

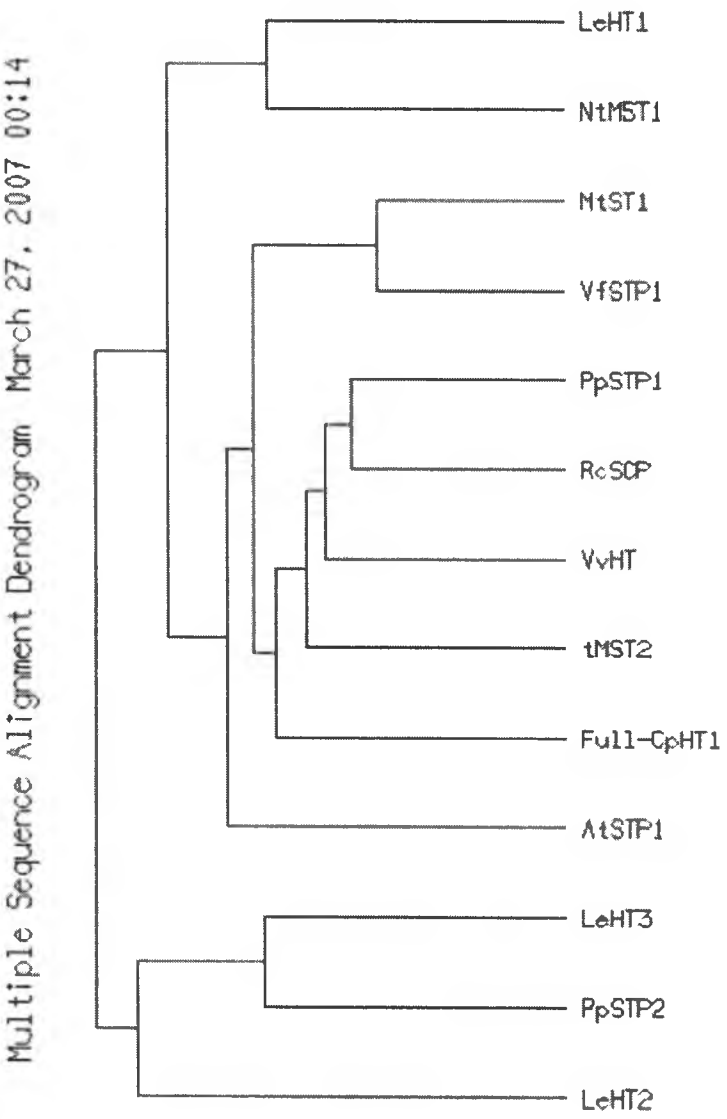

Figure 5.17. (continue).

5.3.7 Total number of hexose transporters in papaya genome

At least 7 possible papaya hexose transporters based upon a search of the sequences from the papaya genome project (http://cqpbr.hawaii.edu/papaya/) using *Arabidopsis AtSTP1* (CAA39037), grape berry *VvHT* (CAA70777) and *VvHT1* (CAA04511), peach fruit *PpSTP1* (*Prunus persica*, AAL16970), and tomato fruit *LeHT1* (CAB52688) and *LeHT3* (CAB52690). This estimation was based on an expected (E) value of e-20 or lower used to evaluate the match when the *PpSTP1* was searched by TBLASTn against the papaya genome. Fifteen to seventeen potential supercontigs possibly contained hexose transporters proteins when searched with *LeHT1, LeHT3, VvHT, AtSTP1* and *VvHT1,* respectively (Table 5.3). Supercontig_1226 contained the full length *CpHT1* (Figure 5.15). Supercontigs 298, 109, 16, 315, 150 and 13 have a highest potential for containing possible hexose transporters.

5.3.8 Structure of the papaya hexose transporter protein

The papaya hexose transporter (*CpHT1*) full-length cDNA clone potentially encodes a protein of 523 amino acids with estimated molecular weight of 57.48 kDA (Figure 5.18). The molecular weight was derived from the 1997 IUPAC standard atomic weights, assuming pH = 7.0. The structure of papaya hexose transporter protein was predicted using the hidden Markov model for transmembrane protein topology prediction program, TMMOD, available at http://liao.cis.udel.edu/website/servers/TMMOD/ (Kahsay *et al.*, 2005). *CpHT1* has 12 transmembrane helices in two 6-helical domain halves separated by the large loop of 60 amino acids between helix number 6 and 7 (Table 5.4 and Figure 5.19). Both amino (N) and carboxyl (C) terminal ends were predicted to locate in the cytoplastic side of the plasma membrane. Comparison of the papaya hexose transporter peptide sequence showed high homology to the sequences from Grape *VvHT* (CAA70777) and *VvHT1* (CAA04511), tomato *LeHT1* (CAB52688), and *Arabidopsis AtSTP1* (CAA39037) (Figure 5.20).

Table 5.3. Hexose transporters were predicted using tBLASTn to occur on a number of papaya genome supercontigs using hexose transporter genes from *Arabidopsis AtSTP1* (CAA39037), grape berry *VvHT* (*Vitis vinifera*, CAA70777) and *VvHT1* (CAA04511), peach fruit *PpSTP1* (*Prunus persica*, AAL16970), and tomato fruit *LeHT1* (*Lycopersicon esculentum*, CAB52688) and *LeHT3* (CAB52690).

AtSTP1			*VvHT*			*VvHT1*		
Papaya	score (bits)	E Value	Papaya	score (bits)	E Value	Papaya	score (bits)	E Value
Supercontig_298	637	0.0	supercontig_298	634	0.0	supercontig_298	637	0.0
Supercontig_1226	373	e-171	Supercontig_1226	378	e-165	supercontig_1226	381	e-169
Supercontig_109	340	e-135	Supercontig_109	342	e-132	supercontig_109	345	e-133
Supercontig_16	470	e-131	Supercontig_16	454	e-127	supercontig_16	458	e-128
Supercontig_315	227	e-127	Supercontig_150	436	e-121	supercontig_13	441	e-123
Supercontig_13	448	e-125	Supercontig_13	430	e-119	supercontig_150	439	e-122
Supercontig_150	440	e-122	Supercontig_315	230	e-119	supercontig_315	234	e-120
Supercontig_882	349	3e-95	Supercontig_882	334	e-107	supercontig_882	343	e-117
Supercontig_226	306	2e-82	Supercontig_226	307	1e-82	supercontig_226	310	2e-83
Supercontig_497	306	3e-82	Supercontig_497	293	2e-78	supercontig_497	297	1e-79
Supercontig_3601	215	7e-55	Supercontig_3601	209	2e-64	supercontig_3601	209	2e-64
Supercontig_24	189	3e-47	Supercontig_24	166	5e-40	supercontig_24	171	1e-41
Supercontig_25	178	1e-43	Supercontig_1686	160	2e-38	supercontig_1686	171	2e-41
Supercontig_1686	176	4e-43	Supercontig_25	157	2e-37	supercontig_25	164	2e-39
Supercontig_258	99	1e-30	Supercontig_258	100	2e-30	supercontig_258	100	2e-31
Supercontig_21	132	6e-30	Supercontig_21	119	7e-26	supercontig_21	128	1e-28
Supercontig_106	115	1e-24				supercontig_106	100	2e-20
possible HT genes		17	possible HT genes		16	possible HT genes		17

PpSTP1			*LeHT1*			*LeHT3*		
Papaya	score (bits)	E Value	Papaya	score (bits)	E Value	Papaya	score (bits)	E Vaiue
Supercontig_1226	120	1e-27	Supercontig_1226	233	e-115	supercontig_150	286	8e-77
Supercontig_497	98	8e-21	Supercontig_150	349	1e-95	supercontig_298	283	9e-76
Supercontig_13	97	1e-20	Supercontig_298	343	6e-94	supercontig_1226	191	6e-74
Supercontig_150	97	2e-20	Supercontig_13	325	2e-88	supercontig_882	272	1e-72
Supercontig_298	96	3e-20	Supercontig_882	313	7e-85	supercontig_13	256	9e-68
Supercontig_882	96	4e-20	Supercontig_16	294	6e-79	supercontig_16	253	6e-67
Supercontig_315	95	9e-20	Supercontig_109	160	9e-79	supercontig_315	177	7e-65
			Supercontig_315	149	1e-69	supercontig_497	236	1e-61
			Supercontig_497	174	6e-43	supercontig_109	154	6e-60
			Supercontig_226	153	2e-36	supercontig_226	165	2e-40
			Supercontig_25	115	3e-25	supercontig_106	110	8e-24
			Supercontig_106	113	2e-24	supercontig_24	108	2e-23
			Supercontig_1686	112	3e-24	supercontig_21	103	7e-22
			Supercontig_24	109	2e-23	supercontig_1686	102	2e-21
			Supercontig_21	100	1e-20	supercontig_25	97	7e-20
possible HT genes		7	possible HT genes		15	possible HT genes		15

Figure 5.18. The ORF of the papaya hexose transporter (*CpHT1*) of 1,572 bp long encoding 523 amino acids. The asterisk (*) showed the terminal codon. The first 'atg' codon is the start codon. The 12 predicted transmembrane helices are underlined.

```
   1 atgcctgcaccaggaggaattgcgccggctgagcccggcagggaa
     M  P  A  P  G  G  I  A  P  A  E  P  G  R  E   15
  46 taccccggtaatcttaccccattcgtcactgtaacatgtatcgtc
     Y  P  G  N  L  T  P  F  V  T  V  T  C  I  V   30
  91 gccgccatgggtggactgatctttggatacgatattgggatctca
     A  A  M  G  G  L  I  F  G  Y  D  I  G  I  S   45
 136 gggggagtgacgtcaatgaactcgtttctgaaggaatttttcccg
     G  G  V  T  S  M  N  S  F  L  K  E  F  F  P   60
 181 gcggttttccggaaaaaggaagaggtatcgtcgactaaccagtac
     A  V  F  R  K  K  E  E  V  S  S  T  N  Q  Y   75
 226 tgtcagtacgacagtccgacacttacgttgtttacatcatcgctg
     C  Q  Y  D  S  P  T  L  T  L  F  T  S  S  L   90
 271 tatctggcggcgcttgtggcgtcgctggttgcggcgacggtgaca
     Y  L  A  A  L  V  A  S  L  V  A  A  T  V  T  105
 316 agaaagttcggtcggaaactgtcgatgctgtttggcggcgtcctg
     R  K  F  G  R  K  L  S  M  L  F  G  G  V  L  120
 361 ttctgcgccggtgccatcattaatggcttcgctaaagctgtttgg
     F  C  A  G  A  I  I  N  G  F  A  K  A  V  W  135
 406 atgttgattctcggcagaattttgttgggttttggcatcggtttt
     M  L  I  L  G  R  I  L  L  G  F  G  I  G  F  150
 451 gccaatcagtctgtaccactctacctctctgagatggctccttac
     A  N  Q  S  V  P  L  Y  L  S  E  M  A  P  Y  165
 496 agatatagaggagcattaaacattggattccaattgtccatcaca
     R  Y  R  G  A  L  N  I  G  F  Q  L  S  I  T  180
 541 attggtattcttgttgccaatgtattgaatttcttctttgcaaaa
     I  G  I  L  V  A  N  V  L  N  F  F  F  A  K  195
 586 atcaaaggaggttggggatggagactgagcttaggaggtgcagta
     I  K  G  G  W  G  W  R  L  S  L  G  G  A  V  210
 631 gttccagctctaatcatcgccattggatcgttaatcctccccgat
     V  P  A  L  I  I  A  I  G  S  L  I  L  P  D  225
 676 acacccaactccatgatcgaacgaggccaggtagaagcagctaaa
     T  P  N  S  M  I  E  R  G  Q  V  E  A  A  K  240
 721 gagaaattaaggagaattcggggtgtcaacaacgtggacgaagag
     E  K  L  R  R  I  R  G  V  N  N  V  D  E  E  255
 766 ttaaaagacttagttgcagcaagtgaagcttcgaaattggtagaa
     L  K  D  L  V  A  A  S  E  A  S  K  L  V  E  270
 811 catccatggagaaacttgttacaaagaaaatacaggcctcatctc
     H  P  W  R  N  L  L  Q  R  K  Y  R  P  H  L  285
 856 accatggctatcatgatcccattcttccagcagctaactggaatt
     T  M  A  I  M  I  P  F  F  Q  Q  L  T  G  I  300
 901 aatgtcatcatgttttatgctcctgttttattcaacacaattggg
     N  V  I  M  F  Y  A  P  V  L  F  N  T  I  G  315
```

```
 946 tttggcagtgacgcctctctcatgtctgctgtaattaccggaatt
     F  G  S  D  A  S  L  M  S  A  V  I  T  G  I   330
 991 gtaaatgtcggtgcaactttggtttcaatctatggagtcgataaa
     V  N  V  G  A  T  L  V  S  I  Y  G  V  D  K   345
1036 tggggaagacgattccttttcctcgagggaggagttcaaatgtta
     W  G  R  R  F  L  F  L  E  G  G  V  Q  M  L   360
1081 atatgccagattgtggtagcagcctccattggagctaaatttggg
     I  C  Q  I  V  V  A  A  S  I  G  A  K  F  G   375
1126 atcaatggcaaccctggagatttaccaaaatggtatgcaattgta
     I  N  G  N  P  G  D  L  P  K  W  Y  A  I  V   390
1171 gtggtgctattcatctgtatttacgtggccggatttgcgtggtct
     V  V  L  F  I  C  I  Y  V  A  G  F  A  W  S   405
1216 tgggggcctctcgggtggctcgtgccgagtgaaatcttccctctt
     W  G  P  L  G  W  L  V  P  S  E  I  F  P  L   420
1261 gaaatcagatcagcagctcagagtatcaatgtgtcggtgaatatg
     E  I  R  S  A  A  Q  S  I  N  V  S  V  N  M   435
1306 ttcttcacatttatagtggcacaaatatttttgacaatgctttgt
     F  F  T  F  I  V  A  Q  I  F  L  T  M  L  C   450
1351 catttgaagtttggactgttcatttttctttgccttctttgtggtt
     H  L  K  F  G  L  F  I  F  F  A  F  F  V  V   465
1396 atcatgtcaatcttcatctactatttcttgccggagacaaaggga
     I  M  S  I  F  I  Y  Y  F  L  P  E  T  K  G   480
1441 atccccattgaagaaatgagtaaggtctggaagtctcactggttc
     I  P  I  E  E  M  S  K  V  W  K  S  H  W  F   495
1486 tggtccaggtttgtggaagatgatggttatggccatcatggaaat
     W  S  R  F  V  E  D  D  G  Y  G  H  H  G  N   510
1531 cttgagatgggaaaaggcaacctgggaccaaagaatgtgtgattc
     L  E  M  G  K  G  N  L  G  P  K  N  V  *      523
1576 atttttctttttttttttttgttttttctgggttttaaagaa

1621 tcggaagcagtaaatagacttttagtttgtttgtattacggaat

1666 tacgaaaaactctcgataatacagatcaattctattcaataatca

1711 tagtaagttatgttctaaaaaaaaaaaaaaaaaaaaaaaaaaa

1756 aaaa
```

Figure 5.18.(continue)

Table 5.4. The 12 transmembrane helices (TMHs) of *CpHT1* protein were predicted using the TMMOD program, http://liao.cis.udel.edu/website/servers/TMMOD/ on March 6, 2007 (Kahsay *et al.*, 2005).

Name	*CpHT1*	
Annotation	TM PROTEIN	
Length	523	
Number of predicted TMHs	12	
Expected number of AAs in TMHs	232.924103	
Expected number, first 60 AAs	18.574062	
Total prob of N-in	0.945778	
Location	Amino acid #	Amino acids
Inside	1 - 23	MPAPGGIAPAEPGREYPGNLTPF
TMhelix 1	24 - 44	VTVTCIVAAMGGLIFGYDIGI
Outside	45 - 82	SGGVTSMNSFLKEFFPAVFRKKEEV SSTNQYCQYDSPT
TMhelix 2	83 - 103	LTLFTSSLYLAALVASLVAAT
Inside	104 - 114	VTRKFGRKLSM
TMhelix 3	115 - 135	LFGGVLFCAGAIINGFAKAVW
Outside	136 - 136	M
TMhelix 4	137 - 157	LILGRILLGFGIGFANQSVPL
Inside	158 - 168	YLSEMAPYRYR
TMhelix 5	169 - 189	GALNIGFQLSITIGILVANVL
Outside	190 - 203	NFFFAKIKGGWGWR
TMhelix 6	204 - 224	LSLGGAVVPALIIAIGSLILP
Inside	225 - 284	DTPNSMIERGQVEAAKEKLRRIRGV NNVDEELKDLVAASEASKLVEHPWR NLLQRKYRPH
TMhelix 7	285 - 305	LTMAIMIPFFQQLTGINVIMF
Outside	306 - 319	YAPVLFNTIGFGSD
TMhelix 8	320 - 343	ASLMSAVITGIVNVGATLVSIYGV
Inside	344 - 355	DKWGRRFLFLEG
TMhelix 9	356 -376	GVQMLICQIVVAASIGAKFGI
Outside	377 - 387	NGNPGDLPKWY
TMhelix 10	388 - 413	AIVVVLFICIYVAGFAWSWGPLGWLV
Inside	414 - 428	PSEIFPLEIRSAAQS
TMhelix 11	429 - 449	INVSVNMFFTFIVAQIFLTML
Outside	450 - 455	CHLKFG
TMhelix 12	456 - 476	LFIFFAFFVVIMSIFIYYFLP
Inside	477 - 523	ETKGIPIEEMSKVWKSHWFWSRFVEDD GYGHHGNLEMGKGNLGPKNV

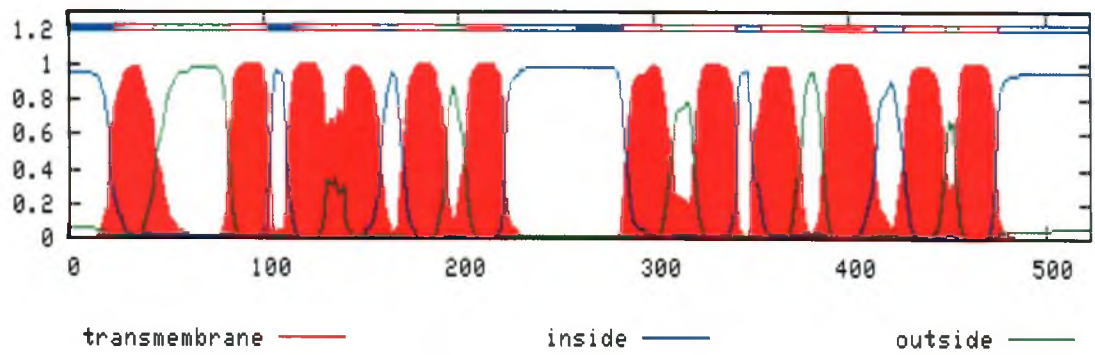

Figure 5.19. The predicted 12 transmembrane helices (TMHs) of the *CpHT1* protein using the TMMOD program available at http://liao.cis.udel.edu/website/servers/TMMOD/ on March 6, 2007 (Kahsay *et al.*, 2005).

Figure 5.20. Multiple comparisons between the full-length papaya hexose transporter peptide sequence (*CpHT1*) and the sequences of grape berry *VvHT* (CAA70777) and *VvHT1* (CAA04511), tomato fruit *LeHT1* (CAB52688) and *Arabidopsis AtSTP1* (CAA39037). The 12 predicted transmembrane helices were underlined. Identical residues conserved among plant hexose transporters were displayed. The sequence alignment was performed using PILEUP from the Pacific Biosciences Research Center, University of Hawaii at Manoa, available at http://www.pbrc.hawaii.edu/ on March 27, 2007.

```
          1                                                  50
 VvHT1pept MPAVGGFDKG T.GKAYPGNL TPYVTVTCVV AAMGGLIFGY DIGISGGVTS
  VvHTpept MPAVGGFDKG T.GKAYPGNL TPYVTVTCVV AAMGGLIFGY DIGISGGVTS
AtSTP1pept MPA.GGFVVG DGQKAYPGKL TPFVLFTCVV AAMGGLIFGY DIGISGGVTS
 CpHT1pept MPAPGGIAPA EPGREYPGNL TPFVTVTCIV AAMGGLIFGY DIGISGGVTS
 LeHT1pept ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  Conserve MPA-GG---- -----YPG-L TP-V--TC-V AAMGGLIFGY DIGISGGVTS
                                             TM1

          51                                                 100
 VvHT1pept MAPFLQKFFP SVYRKEALDK STNQYCKFDS ETLTLFTSSL YLAALLSSLV
  VvHTpept MAPFLQKFFP SVYRKEALDK STNQYCKFDS ETLTLFTSSL YLAALLSSLV
AtSTP1pept MPSFLKRFFP SVYRKQQEDA STNQYCQYDS PTLTMFTSSL YLAALISSLV
 CpHT1pept MNSFLKEFFP AVFRKKEEVS STNQYCQYDS PTLTLFTSSL YLAALVASLV
 LeHT1pept ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  Conserve M--FL--FFP -V-RK----- STNQYC--DS -TLT-FTSSL YLAAL--SLV
                                                    TM2

          101                                                150
 VvHT1pept AATVTRKFGR KLSMLFGGLL FCAGAIINGA AKAVWMLIVG RILLGFGIGF
  VvHTpept AATVTRKFGR KLSMLFGGLL FCAGAIINGA AKAVWMLIVG RILLGFGIGF
AtSTP1pept ASTVTRKFGR RLSMLFGGIL FCAGALINGF AKHVWMLIVG RILLGFGIGF
 CpHT1pept AATVTRKFGR KLSMLFGGVL FCAGAIINGF AKAVWMLILG RILLGFGIGF
 LeHT1pept ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  Conserve A-TVTRKFGR -LSMLFGG-L FCAGA-ING- AK-VWMLI-G RILLGFGIGF
                                  TM3                 TM4

          151                                                200
 VvHT1pept ANQSVPLYLS EMAPYKYRGA LNIGFQLSIT IGILVANILN YFFAKIKGGW
  VvHTpept ANQSVPLYLS EMAPYKYRGA LNIGFQLSIT IGILVANILN YFFAKIKGGW
AtSTP1pept ANQAVPLYLS EMAPYKYRGA LNIGFQLSIT IGILVAEVLN YFFAKIKGGW
 CpHT1pept ANQSVPLYLS EMAPYRYRGA LNIGFQLSIT IGILVANVLN FFFAKIKGGW
 LeHT1pept ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~
  Conserve ANQ-VPLYLS EMAPYKYRGA LNIGFQLSIT IGILVA--LN -FFAKIKGGW
                                        TM5

          201                                                250
 VvHT1pept GWRLSLGGAV VPALIITVGS LVLPDTPNSM IERGQHEGAK TKLRRIRGVD
  VvHTpept GWRLSLGGAV VPALIITVGS LVLPDTPNSM IERGQHEGAK TKLRRIRGVD
AtSTP1pept GWRLSLGGAV VPALIITIGS LVLPDTPNSM IERGQHEEAK TKLRRIRGVD
 CpHT1pept GWRLSLGGAV VPALIIAIGS LILPDTPNSM IERGQVEAAK EKLRRIRGVN
 LeHT1pept ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~NHDEAK ARLKRIRGIE
  Conserve GWRLSLGGAV VPALII--GS L-LPDTPNSM IERG----AK --L-RIRG--
                     TM6
```

```
          251                                                       300
 VvHT1pept DVEEEFNDLV VASEASKLVE HPWRNLLQRK YRPHLTMAIL IPFFQQLTGI
  VvHTpept DVEEEFNDLV VASEASKLVE HPWRNLLQRK YRPHLTMAIL IPFFQQLTGI
AtSTP1pept DVSQEFDDLV AASKESQSIE HPWRNLLRRK YRPHLTMAVM IPFFQQLTGI
 CpHT1pept NVDEELKDLV AASEASKLVE HPWRNLLQRK YRPHLTMAIM IPFFQQLTGI
 LeHT1pept DVDEEFNDLV IASEASRKIE HPWRNLLQKK YRPHLTMAIM IPFFQQLTGI
  Conserve -V--E--DLV -AS--S---E HPWRNLL--K YRPHLTMA-- IPFFQQLTGI
                                                        TM7

        301                                                       350
 VvHT1pept NVIMFYAPVL FKTIGFADDA SLMSAVITGG VNVLATIVSI YGVDKWGRRF
  VvHTpept NVIMFYAPVL FKTIGFADDA SLMSAVITGR VNVLATIVSI YGVDKWVRRF
AtSTP1pept NVIMFYAPVL FNTIGFTTDA SLMSAVVTGS VNVGATLVSI YGVDRWGRRF
 CpHT1pept NVIMFYAPVL FNTIGFGSDA SLMSAVITGI VNVGATLVSI YGVDKWGRRF
 LeHT1pept NVIMFYAPVL FKTIGFGTDA SLMSAVITGG INVIATIVSI YYVDKLGRRF
  Conserve NVIMFYAPVL F-TIGF--DA SLMSAV-TG- -NV-AT-VSI Y-VDK--RRF
                                          TM8

          351                                                     400
 VvHT1pept LFLEGGTQML ICQVIVATCI GVKFGVDGEP GALPKWYAIV VVLFICVYVS
  VvHTpept LFLEGGTQML ICQVIVATCI LVKFGVDGEP WCLPKWYAIV VVLFICVYVS
AtSTP1pept LFLEGGTQML ICQAVVAACI GAKFGVDGTP GELPKWYAIV VVTFICIYVA
 CpHT1pept LFLEGGVQML ICQIVVAASI GAKFGINGNP GDLPKWYAIV VVLFICIYVA
 LeHT1pept LFLEGGIQML FSQIAVAILI AIKFGVNGTP GELPKWYAIV VVIFICVYVA
  Conserve LFLEGG-QML --Q--VA--I --KFG--G-P --LPKWYAIV VV-FIC-YV-
                     TM9                                  TM10

          401                                                     450
 VvHT1pept GFAWSWGPLG WLVPSEIFPL EIRSAAQSVN VSVNMFFTFI IAQIFLNMLC
  VvHTpept GFAWSWGPLG WLVPSEIFPL EIRSAAQSVN VSVNMFFTFI IAQIFLNMLC
AtSTP1pept GFAWSWGPLG WLVPSEIFPL EIRSAAQSIT VSVNMIFTFI IAQIFLTMLC
 CpHT1pept GFAWSWGPLG WLVPSEIFPL EIRSAAQSIN VSVNMFFTFI VAQIFLTMLC
 LeHT1pept GFAWSWGPLG WLVPSEIFPL EIRSAAQSIN VSVNMIFTFA VAQVFLTMLC
  Conserve GFAWSWGPLG WLVPSEIFPL EIRSAAQS-- VSVNM-FTF- -AQ-FL-MLC
                                                  TM11

          451                                                     500
 VvHT1pept HMKFGLFLFF AFFVVVMSFF IYFFLPETKG IPIEEMAEVW KSHWFWSRYV
  VvHTpept HMKFGLFLSF AFFVVVMSFF IYFFLPETKG IPIEEMAEVW KSHWFWSRYV
AtSTP1pept HLKFGLFLVF AFFVVVMSIF VYIFLPETKG IPIEEMGQVW RSHWYWSRFV
 CpHT1pept HLKFGLFIFF AFFVVIMSIF IYYFLPETKG IPIEEMSKVW KSHWFWSRFV
 LeHT1pept HLKFGLFLFF AFFVVIMTVF IYFFLPETKN IPIEEMVIVW KEHWFWSKFM
  Conserve H-KFGLFL-F AFFVV-M--F -Y-FLPETK- IPIEEM--VW --HW-WSR--
                     TM12

          501                              528
 VvHT1pept NDGSY..... .SGVELVKEN YPV..KNV
  VvHTpept NDGSY..... .SGVELVKEN YPV..KNV
AtSTP1pept EDGEY...GN ..ALEMGKNS NQAGTKHV
 CpHT1pept EDDGY...GH HGNLEMGKGN ..LGPKNV
 LeHT1pept TEVDYPGTRN GTAVEMAKGG ..AGYKIV
  Conserve ----Y----- ----EM-K-- -----K-V
```

Figure 5.20. (continue).

Four long identical conserved regions were found between Val-30 and Met-51, Val-155 and Ala-186, Phe-192 and Ile-216, and Gly-401 and Ser-428. *CpHT1* contained three potential *N*-glycosylation sites at N (Asn) residue 19, 152 and 430 since they belong to the Asn-X-Ser or Asn-X-Thr consensus sequences (Figure 5.20). Asn-19 is located in the cytosol, whereas Asn-152 and Asn-430 are located in the fourth and eleventh transmembrane helices, respectively (Table 5.4). Therefore, *CpHT1* is not likely to be glycosylated except at site Asn-19.

A search of N-terminus signal within *CpHT1* peptide sequence using SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/) indicated that the papaya hexose transporter protein does not contain an N-terminal signal sequence.

5.4 Discussion

Sweetness of papaya fruit depends on the accumulation of sugar during fruit development, maturation and ripening. Papaya fruit accumulates sugar in the form of sucrose (Chan *et al.*, 1979). Since no starch degradation appears to occur, papaya fruit sweetness relies on its sugar content at the time of harvest. Zhou and Paull (2001) proposed that sugar accumulation during fruit maturation and ripening involved apoplastic phloem unloading due to the high activity of the cell wall acid invertase at 125 DAA. Acid invertase inverts sucrose into the two hexose sugars, glucose and fructose. The discovery of the first papaya hexose transporter (*CpHT1*) supports the apoplastic unloading in papaya fruit during maturation and ripening. Hexose transporters in fruit tissues have been reported in grape (Fillion *et al.*, 1999; Vignault *et al.*, 2005), tomato (Gear *et al.*, 2000) and peach (Etienne *et al.*, 2002).

Southern blot analysis confirmed the presence of a gene family of hexose transporters in tomato consisting of at least three members. The three, *LeHT1, LeHT2* and *LeHT3*, were cloned (Gear *et al.*, 2000). *LeHT2* is highly expressed in source leaves and certain sink tissues such as flowers. *LeHT1* and *LeHT3* are highly expressed in young fruit (10 DAA) and then expression declines until 24 DAA (Gear *et al.*, 2000). *LeHT3* expression is detectable through out the ripening stage. *LeHT2* is expressed in tomato fruit at the onset of fruit ripening (Dibley *et al.*, 2005).

In papaya, a search of the hexose transporter cDNA (*CpHT1*) against the papaya genomic DNA database indicated that *CpHT1* identical to a sequence on supercontig-1226. The number of papaya hexose transporters was estimated to be between seven and seventeen when searched with the peach *PpSTP1* and grape *VvHT1* sequences (Table 5.3). This number falls within the reported range of seven hexose transporter genes found in grape and fourteen in *Arabidopsis*.

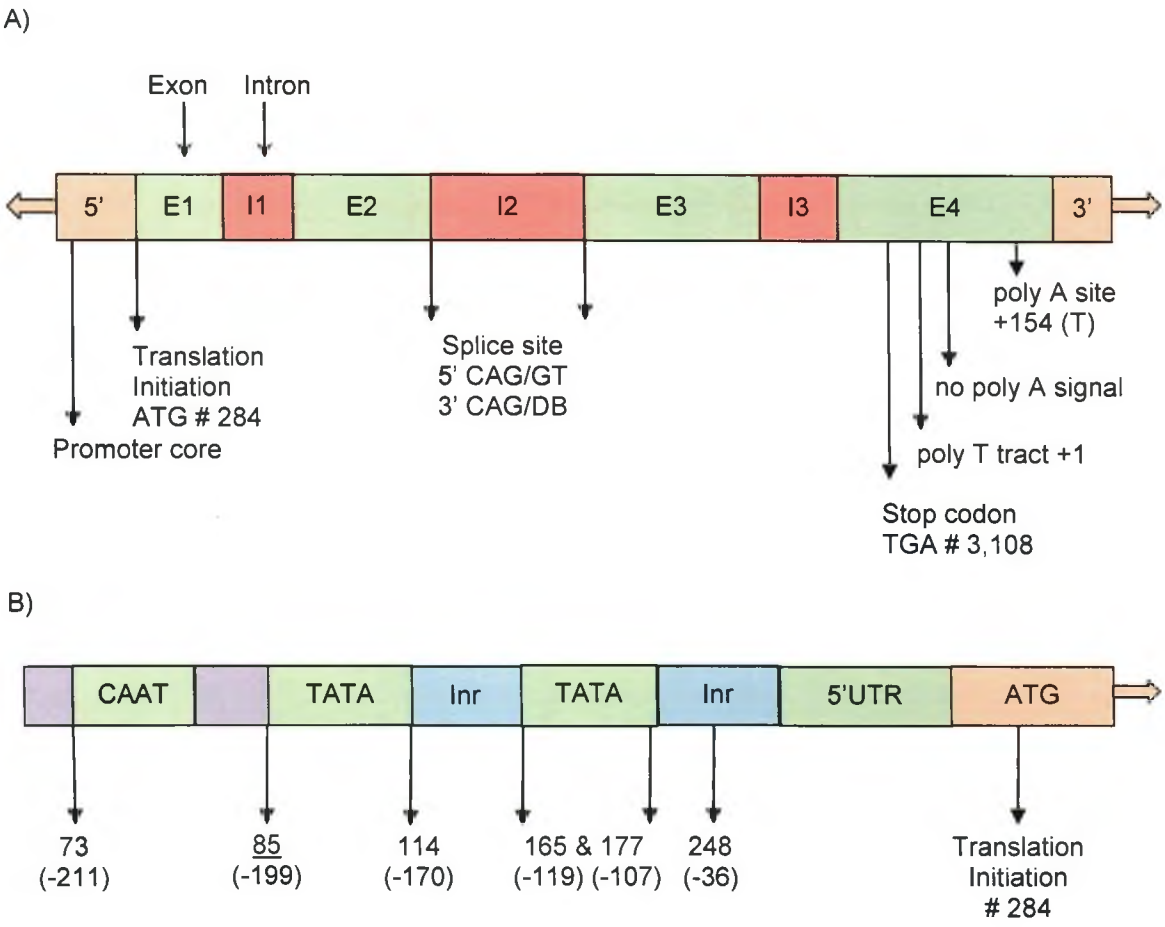

Figure 5.21 A) *CpHT1* possible gene structure, B) *CpHT1* possible promoter structure

The sequence of papaya hexose transporter cDNA (*CpHT1*) amplified from fruit tissue was identical to the papaya genomic DNA sequence on supercontig-1226 and had a sequence similar to that of grape *VvHT1* that was amplified from fruit berry (Fillion *et al.*, 1999). The predicted *CpHT1* gene structure was similar to other hexose transporters. *CpHT1* contains four exons and three introns similar to *VvHT* (Figure 5.15 and 5.21 A). The translation initiation codon (ATG) is located at 284 nt of the genomic *CpHT1* sequence. This ATG more likely represents the functional translation start codon, since it gave the longest open reading frame of 523 amino acids (Figure 5.14). Similar to *VvHT1*, the ATG of *CpHT1* was not located in either [C/G]AANNATGG or TAAACAATGGCT consensus sequences that have been described for the translation initiation site in plants (Joshi, 1987; Lu¨tcke *et al.*, 1987). The sequence between the site of translation initiation codon and the poly A site (T) was 2,979 bp long, 158 nt longer than the *VvHT1* at 2,821 bp (Fillion *et al.*, 1999).

The 5' untranslated region (5'UTR) is essential for translation and affects RNA stability (Schuster *et al.*, 1999). Since none of the cDNA clones amplified in this study contained the 5'UTR region with the start codon, the size of the 5'UTR and the location of the 5'UTR start point of *CpHT1* were unclear. The sequences of the *CpHT1* cDNA clones was identical to the genomic sequence, therefore the 5'UTR of the genomic *CpHT1* DNA contained one potential CCAAT and three possible TATA motifs. The CCAAT motif started at – 211 nt upstream from the translation initiation codon (ATG) (Figure 5.15 and 5.21 B). Three possible TATA motifs were located at – 199, – 119 and – 107 nt upstream from the start point of the *CpHT1*. Unlike *VvHT1* (Fillion *et al.*, 1999) and other genes in many plants (Hanley and Schuler, 1988), none of the *CpHT1* TATA boxes corresponded to the consensus sequence (C/G)TATA(T/A)A1–3(C/T)A for TATA boxes. This core promoter consisting of the CCAAT and TATA motifs is necessary for RNA polymerase II (Becker *et al.*, 2000). The initiator site (Inr), a pyrimidine rich region, was located between -36 and -170 covering TATA-119 and TATA-107 boxes. The location and length of the pyrimidine rich region suggested that the TATA-119 and TATA-107 boxes were probably not functional and were parts of the initiator site. *CpHT1* contained 35-nt long polypurine (A) - rich sequence (PARS) between -1 and -35 upstream from the start point. This papaya PARS region

also consisted of five repeated [GAAA] sequences. The appearance of 35-nt long PARS element with $[GAAA]_5$ in the 5'UTR of *CpHT1* suggested that the translation of *CpHT1* was possibly an IRES-dependent mechanism instead of the traditional cap-dependent mechanism (Dorokhov *et al.*, 2002). Two long (32- and 34-nt) and several short PARS elements with multiple (G)1–4(A)2–5 modules are also found in the 453-nt 5'UTR of *N. tabacum* heat-shock factor 1 mRNA. This long PARS naturally occurs in long 5'UTR of plant mRNAs and promotes the cross-kingdom conservation of internal ribosome entry site (IRES) activity (Dorokhov *et al.*, 2002). The cap-independent translation mechanism has been reported in plant viral RNAs (Kneller *et al.*, 2006). The polypurine rich region in the 5'UTR of the *CpHT1* also has the same consensus sequence of the GA binding site (RGARAGRRA) found in *Arabidopsis* gene SEEDSTICK (*STK*), an homeotic gene that controls ovule identity and is characterized by its mechanism (Kooiker *et al.*, 2005). Twelve purine-rich elements presented in the regulatory sequence within the 5'UTR of *Arabidopsis STK* were regulated by the GA binding protein (GBP), BASIC PENTACYSTEINE1 (BPC1). The binding of BPC1 to the purine rich element induces conformational changes in the *STK* regulatory region (Kooiker *et al.*, 2005).

Unlike the *VvHT1* promoter (Fillion *et al.*, 1999), none of the repetitive A/T sequences, SEF4 motif (RTTTTTR), CAGAAGATA, and TGTAAAG consensus, which are characteristic features of plant promoters, were found in the 5'UTR of *CpHT1*. However, the repetitive sequence CANNTG that correspond to the E-box/ABA-responsive elements (Stalberg *et al.*, 1996) was frequently found in the 5'UTR, introns and the coding region of *CpHT1* as in *VvHT1*.

The genomic *CpHT1* introns contain T(U)A-rich sequences which are the main distinguishing characteristic of plant introns. Miller *et al.* (2001) reported that UA-rich sequences in plants likely acted early in intron recognition and definition, and directed the association of splicing factors to assemble presplicing complexes. The binding of the first U-rich RNA binding protein, UBD-1, to the UA-rich region increases the efficiency of splicing of weak introns and also increases transcript levels. The splicing consensus sequences AG/GT (Breathnach and Chambon, 1981; Hanley and Schuler, 1988) were conserved at the 5' splice site of the introns of

CpHT1 and in *VvHT1*. In addition, CAG codon was observed on both 5' and 3' splice sites of each introns (Figure 5.15, 5.16, and 5.21 A). The location of these CAG codons that cover the entire intron of each hexose transporter, were also found in grape full length *VvHT1* cDNA. The MWVYCAG/GTWTD consensus acts as a 5' recognition splice site and the WWWVCA/GDBKGW consensus acts as a 3' recognition splice site for *CpHT1*. Similarly to many plants introns (Brown, 1986), the 5' splice site of *CpHT1* was similar to the one in the mammalian consensus, CAGGTAAGT (Mount, 1982).

The 3' untranslated end of the cDNA contained a poly-T tract, 1 nt downstream from the stop codon (TGA) (Figure 5.15 and 5.21 A). This poly T tract is also found in *VvHT1* (AJ001061), however, it is located at 147 nt downstream from the stop codon. This short sequence of U's is a transcriptional termination signal necessary for RNA polymerase III which is mostly found in prokaryotes (Becker *et al.*, 2000). The possible transcription termination signal, AAT(U)AAA, in *CpHT1* was found 131 nt downstream from the termination codon (TGA) and -24 nt upstream from the poly A site (T). However, this signal (AATAA) was incomplete as it lacked one last adenylate (A). This short sequence is probably not the termination signal. Other AATAAAs were found 3' to the poly A site (T) (Figure 5.15). Therefore, *CpHT1* probably does not contain the eukaryotic polyadenylation signal (AATAAA) at the 3' untranslated end. This is the same as with the three *AtSTP1* clones (Sauer *et al.*, 1990b) and grape *VvHT1* (Fillion *et al.*, 1999). The lack of AATAAA in *CpHT1*, *AtSTP1* and *VvHT1*, combined with the appearance of poly T tract following the stop codon, suggest that transcription termination and polyadenylation of plant hexose transporter genes are possibly the result of RNA polymerase III.

The papaya hexose transporter (*CpHT1*) full-length cDNA encodes a protein of 523 amino acids, with a calculated molecular mass of 57.48 kDa (Figure 5.14). A search of papaya hexose transporter peptide sequences against BLASTX showed the five peptide sequences with the highest homology were poplar *tMST2.2* (94% similarity; CAG27609), English walnut *JrHT1* (94% similarity, AAY89231), *Datisca glomerata DtHT1* (92% similarity; CAD30830), castor bean *RcSTC* (93% similarity; Q41144), and *Arabidopsis AtSTP1* (83% similarity; CAA39037). Grape *VvHT* had 91% similarity (CAA04511 and CAA70777), Olive *OeMST2* 81% similarity (ABJ98314),

tomato *LeHT2* 76% (CAB52689) and *LeHT1* 85% (CAB52688), and rice *OsMST3* 81% (NP_001059981) and *OsMST6* 79% (AAQ24872).

The structure of the papaya hexose transporter protein was predicted from the encoded 523 amino acid sequence using the hidden Markov model for transmembrane protein topology prediction program, TMMOD, available at http://liao.cis.udel.edu/website/servers/TMMOD/ (Kahsay *et al.*, 2005). TMMOD was developed from the prototype TMHMM, a transmembrane protein topology prediction method available at http://www.cbs.dtu.dk/services/TMHMM/ (Krogh *et al.*, 2001), but differed from TMHMM by the architecture of the submodels for loops on both sides of the membrane and the model training procedure. Prediction of 83 transmembrane proteins with known topology showed that TMMOD outperformed TMHMM and other existing methods with an accuracy of 89% for both topology and locations (Kahsay *et al.*, 2005). The prediction showed that *CpHT1* had 12 transmembrane helices in two 6-helical domain halves separated by a large loop of 60 amino acids between helix number 6 and 7 (Table 5.3). The deletion of amino acid residues in this large cytoplasmic loop of the human glucose transporter *Glut1* progressively reduces the uptake of the 2-deoxy-D-glucose by *Glut1* expressed in *Xenopus oocytes*. Hence, this large cytoplasmic loop is essential for its proper function (Monden *et al.*, 2001). Both the N-terminus and C-terminus of *CpHT1* were predicted to locate in the cytoplasmic side of the plasma membrane. A search of N-terminus signal within *CpHT1* peptide sequence using SignalP 3.0 server (http://www.cbs.dtu.dk/services/SignalP/) indicated that the papaya hexose transporter protein does not contain an N-terminal signal sequence and has no site for N-glycosylation similar to *Chlorella* hexose transporter (Sauer and Tanner, 1989), *AtSTP1*, or other homologous hexose transporter proteins (Sauer *et al.*, 1990b). *CpHT1* contains a LPETK motif (position 475–479) in the C-terminal part at the beginning of the last hydrophilic loop. This sequence was immediately followed by GIPIEEMSKVWKSHWFWSR, which matched the conserved consensus (M/V)XX(V/L)(W/Y)XXHW(F/Y)WX(R/K). This highly conserved motif and consensus are characteristic features of all hexose transporters cloned so far from higher plants (Fillion *et al.*, 1999).

The structure of the *CpHT1* polypeptide was similar to hexose transporter polypeptide of other plants including tomato *LeHT2* (Gear *et al.*, 2000). The full-length *LeHT2* cDNA also encodes a protein of 523 amino acids, with a calculated molecular mass of 57.6 kDa. The *Arabidopsis* hexose transporter protein (*AtSTP1*) has a length of 522 amino acids yielding a calculated molecular mass of 57.5 kDa (Sauer *et al.*, 1990b). The predicted *LeHT2* protein has 12 putative membrane-spanning domains and belongs to the Major Facilitator Superfamily of membrane carriers (Gear *et al.*, 2000).

The functions of each transmembrane helix and loop are still unclear. Hruz and Mueckler (2001) suggested that the inner and outer glucose-binding sites of human glucose transporter 1 (GLUT1) were probably located in transmembrane segments 9, 10, 11 and the QLS motif located in the seventh transmembrane segment could be involved in the selection and affinity of transported substrate (Seatter *et al.*, 1998; Hruz and Mueckler, 1999).

5.5 Summary

The first papaya hexose transporter cDNA (*CpHT1*) was cloned from 'Sunset' papaya fruit. The sequence of *CpHT1* cDNA matched part of the sequence on the papaya genome supercontig_1226. The *CpHT1* DNA found in supercontig_1226 consists of four exons and three introns. The 5' untranslated region contains a possible CCAAT motif located at – 211 nt upstream from translation initiation codon (ATG) and three possible TATA motifs located at – 199, – 119 and – 107 nt upstream from ATG. The initiator site (Inr), pyrimidine rich region located between the CCAAT and the 35-nt-purine rich region covering the TATA motif. The 3' untranslated region contains a short sequence of poly T(U) tract, 1 nt downstream from the stop codon (TGA). The possible transcription termination signal, AAT(U)AAA was found at 131 nt downstream from the termination codon (TGA) or -24 nt upstream from the poly A site (T). The total number of hexose transporters in papaya was estimated to be at least seven.

The full length 1,732 bp *CpHT1* mRNA encodes a 523 amino acid protein. The protein was estimated to be 57.48 kDa and contain 12 transmembrane helices with both amino and carboxyl terminals located in the cytosol. The newly characterized papaya hexose transporter

probably does not contain an N-glycosylation site.

CHAPTER 6

SUMMARY

'Sunset' papaya had smaller fruit size but contained higher total soluble solids, total sugar, dry weight and total protein than those of UH801, a low sugar breeding line. Maturation and sugar accumulation of 'Sunset' occurred two to three weeks prior to its occurrence in UH801. Papaya hexose transporters appear to be an energy-dependent cotransporter. However, the activity of hexose transporter of 'Sunset' fruit was less than those of UH801.

The first papaya hexose transporter cDNA (*CpHT1*) was cloned from 'Sunset' papaya fruit. The *CpHT1* gene was identical in papaya genome supercontig_1226. The *CpHT1* DNA sequence found in supercontig_1226 predicts four exons and three introns. The 5' untranslated region contains a possible CCAAT box located at – 211 nt upstream from transcription start point (TSP). Three possible TATA boxes were located at – 199, – 119 and – 107 nt upstream from TSP. The initiator site (Inr), pyrimidine rich region was located between the CCAAT and TATA boxes and also between the TATA box and TSP. The 3' untranslated region contains a short sequence of poly T(U) tract, 1 nt downstream from the stop codon (TGA). The possible transcription termination signal, AAT(U)AAA was found at 131 nt downstream from the termination codon (TGA) and -24 nt upstream from the poly A site (T). The total number of hexose transporters in papaya was estimated to be at least seven.

The full length 1,732 bp *CpHT1* mRNA encodes a 523 amino acid protein. The protein was estimated to be 57.48 kDA and contained 12 transmembrane helices with both amino and carboxyl terminal located in cytosol.

In future studies, particular attention and resources should be devoted to elucidate *CpHT1* gene expression pattern during fruit maturation and ripening; determine the total number of papaya hexose transporters, their locations and expression; and perform transformation of *CpHT1* gene using fruit specific promoter. Other topics may present themselves with progressing research.

REFERENCES

Akamine, E.K., Goo, T. 1971. Relation between surface color development and total soluble solids in papaya. *HortScience*. 6: 567-568.

Ali, Z.M., Lizada, M.C.C., Lazan, H. 1994. Ripening and senescence. P56-74 In Rohani, M.Y. (ed.). Papaya-Fruit Development, Postharvest Physiology, Handling and Marketing in ASEAN. ASEAN Food Handling Bureau. Malaysia.

Antognozzi, E., Tombesi, A., Palliotti, A. 1992. Relationships between leaf area, leaf area index and fruiting in kiwifruit (*Actinidia deliciosa*). Acta Hort. 297:435-442.

Anon. 1990. Standards for Hawaii-grown papaya. Hawaii Department of Agriculture, Marketing Division.

Awada, M. 1958. Relationship of minimum temperature and growth rate with sex expression of papaya plants (*Carica papaya* L.). Hawaii Agricultural Experiment StationTechnical Bulletin 38, University of Hawaii, Honolulu, *cited by* Nakasone, H.Y. and Pauli, R.E. 1998. Tropical Fruits, Chapter 10 Papaya. P239-269. CAB international.

Awada, M., Ikeda, W.S. 1957. Effects of water and nitrogen application on composition, growth, sugars in fruit, yield, and sex expression of the papaya plants. Hawaii Agric Exp.Stn.Tech.Bull. 33.

Badillo, V.M. 2000. *Carica* L. vs. *Vasconcella* St. Hil. (Caricaceae) con la Rehabilitacion de este Ultimo. Ernstia 10: 74-79.

Becker, W. M., Kleinsmith, L. J, Hardin, J. 2000. The World of the Cell, Fourth edition. The Benjamin/Cummings Publishing Company. San Francisco, CA. 878 p.

Beffagna, N., Romani, G. 1988. Effects of two plasmalemma ATPase inhibitors on H^+ Extrusion and Intracellular pH in *Elodea densa* leaves. Journal of Experimental Botany 39(205):1033-1043.

Boorer, K. J., Loo, D. D. F., Wright, E. M. 1994. Steady-state and presteady-state kinetics of the H^+/hexose cotransporter (STP1) from *Arabidopsis thaliana* expressed in *Xenopus oocytes*. J.

Biol. Chem. 269:20417-20424.

Breathnach, R., Chambon, P. 1981. Organization and expression of split genes coding for proteins. Annu Rev Biochem 50:349–383.

Brown, J. W. S. 1986. A catalogue of splice junction and putative branch point sequences from plant introns. Nucl. Acids Res. 14:9549-9559.

Brown, M. M., Hall, J. L., Ho, L. C .1997. Sugar uptake by protoplasts isolated from tomato fruit tissues during various stages of fruit growth. Physiol. Plant 101: 533–539.

Bruneau, J. M., Worrell, A. C., Camdou, B. 1991. Sucrose phosphate synthase, a key enzyme for sucrose biosynthesis in plants, protein purification from corn leaves and immunological detection. Plant Physiol. 96:473.

Bush, D. R. 1993. Proton-coupled sugar and amino acid transporters in plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. 44:513-542.

Büttner, M., Sauer, N. 2000. Monosaccharide transporters in plants: structure, function and physiology. Biochimica et Biophysica Acta 1465:263-274.

Büttner, M., Truernit, E., Baier, K., Scholz-Starke, J., Sontheim, M., Lauterbach, C., Hub, V. A. R., Sauer, N. 2000. *AtSTP3*, a green leaf-specific, low affinity monosaccharide-H^+ symporter of *Arabidopsis thaliana*. Plant Cell Environ. 23:175-184.

Canel, C., Bailey-Serres, J. N., Roose, M. L. 1995. In vitro [^{14}C] citrate uptake by tonoplast vesicles of acidless Citrus juice cells. J.Amer.Soc.Hort.Sci. 120:510-514.

Chacko, E. K., Reddy, Y. T. N., Annthanaranan, T. V. 1982. Studies on the relationship between leaf number and area and fruit development in mango. J.Hort.Sci. 57:483-492.

Chan, H. T., Kwok, S. C. M. 1976. Isolation and characterization of β-fructofuranosidase from papaya. J. Food Sci. 41:320-323.

Chan, H. T., Kenneth, L. H. Goo, T., Akamine, E. K. 1979. Sugar composition of papaya during fruit development. HortSci. 14:140-141.

Cheikh, N., Brenner, M. L., Huber, J. L., Huber, S. C. 1992. Regulation of sucrose phosphate synthase by gibberellins in soybean and spinach plants. Plant Physiol. 100:1238-1242.

Chen, C. C., Zhou, L. Chen, N. J., Paull, R. E. 2001. Determinants of sweetness in papaya and

pineapple. Acta Hort. 553:99-100.

Cocucci, M. C. 1986. Inhibition of plasmalemma and tonoplast ATPase by erythrosin B. Plant Sci. 47:221-227.

DeJong, T. M., Doyle, J. F., Day, K. R. 1987. Seasonal patterns of reproductive and vegetative sink activity in early and late maturing peach (*Prunus persica*) cultivars. Physiol. Plant. 71:83-88.

Dibley, S. J., Gear, M. L., Yang, X., Rosche, E. G., Offler, C. E., McCurdy, D. W., Patrick, J. W. 2005. Temporal and spatial expression of Hexose Transporters in developing tomato fruit. Functional Plant Biology. 32(9):777–785.

Ding, B., Parthasarathy, M.V., Niklas, K.,Turgeon, R. 1988. A morphometric analysis of the phloem-unloading pathway in developing tobacco leaves. Planta. 176: 307-318.

Dorokhov, Y. L., Skulachev, M. V., Ivanov, P. A., Zvereva, S. D., Tjulkina, L. G., Merits, A., Gleba, Y. Y., Hohn, T., Atabekov, J. G. 2002. Polypurine (A)-rich sequences promote cross-kingdom conservation of internal ribosome entry. PNAS 99:5301-5306.

Doyle, J. J., Doyle, J. L. 1990. Isolation of plant DNA from fresh tissue. Focus 12(1): 13-15.

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., Smith, F. 1956. Colorimetric method for determination of sugars and related substances. Anal. Chem. 28:350-356.

Eleftheriou, E. P. 1987. Changes in the endoplasmic reticulum during sieve element differentiation in *Triticum* aestivum. Ann. Bot. 60:713-722.

Esau, K., Gill, R. H. 1971. Aggregation of endoplasmic retlculum and its relation to the nucleus in a differentiating sieve element. J. Ultrastruct. Res. 34, 144-158.

Esau, K., Gill, R. H. 1972. Nucleus and endoplasmic reticulum in differentiating root protophloem of *Nicotiana tabacum*. J. Ultrastruct. Res. 41, 160-1 75.

Etienne, C., Rothan, C., Moing, A., Plomion, C., Bodenes, C., Svanella-Dumas, L., Cosson, P., Pronier, V., Monet, R., Dirlewanger, E. 2002. Candidate genes and QTLs for sugar and organic acid content in peach [*Prunus persica* (L.) Batsch]. Theor. Appl. Genet. 105:145-159.

Fillion, L., Ageorges, A., Picaud, S., Coutos-Thevenot, P., Lemoine, R., Romieu, C., Delrot, S. 1999. Cloning and expression of a hexose transporter gene expressed during the ripening of

grape berry. *Plant Physiology* 120:1083-1093

Fisher, D. B., Wu, Y., Ku, M. S. B. 1992. Turnover of soluble proteins in the wheat sieve tube. Plant Physiol. 100, 1433-1441.

Fourney, R.M., Miyakoshi, J., Day III, R.S., Paterson, M.C. 1988. Northern Blotting: Efficient RNA staining and transfer. *Focus* 10(1): 5-7

Gear, M. L., McPhillips, M. L., Patrick, J. W., McCurdy, D. W. 2000. Hexose transporters of tomato: molecular cloning, expression analysis and functional characterization. Plant Molecular Biology 44:687-697

Geiger, D. R., Giaquinta, R. T., Sovonick, S. A., Fellows, R. J. 1973. Solute distribution in sugar beet leaves in relation to phloem loading and translocation. Plant Physiol. 52:585–589.

Genard, M., Bertin, N., Borel, C., Bussie`res, P., Gautier, H., Habib, R., Le´chaudel, M., Lecomte, A., Lescourret, F., Lobit, P., Quilot, B. 2007. Towards a virtual fruit focusing on quality, modeling features and potential uses. J. Exp. Bot. 58:917–928.

Getz, H. P. 1991. Sucrose transport in tonoplast vesicles of red beet roots is linked to ATP hydrolysis. Planta 185:261-268.

Godt, D.E., Roitsch, T. 1997. Regulation and tissue-specific distribution of mRNAs for three extracellular invertase isoenzymes of tomato suggests an important function in establishing and maintaining sink metabolism. Plat Physiol. 115:273-282.

Gonslaves, D. 2004.Transgenic papaya in Hawaii and beyond. AgBioForum 7: 36-40.

Gougler Schmalstig, J., Geiger, D. R. 1985. Phloem unloading in developing leaves of sugar beet 1. Evidence for pathway through the symplast. Plant Physiol. 79:237-241.

Griffith, J.K., Baker, M.E., Rouch, D.A., Page, M.G., Skurray, R.A., Paulsen, I.T., Chater, K.F., Baldwin, S.A., Henderson, P.J. 1992. Membrane transport proteins: implications of sequence comparisons. Curr. Opin. Cell Biol. 4:684-695.

Hanley BA, Schuler MA (1988) Plant intron sequence: evidence for distinct groups of introns. Nucleic Acids Res 16: 7159–7175.

Hawaii Agricultural Statistics Service. 2004. Statistics of Hawaii Agriculture 2002. Hawaii Department of Agriculture, Agricultural Development Division. 101 p.

Hosaka, M., Kanayama, Y., Shiratake, K., Yamaki, S. 1994. tonoplast H^+-ATPase of mature pear fruit. Phytochemistry 36:565-567.

Hruz, P.W., Mueckler, M.M. 1999. Cysteine-scanning mutagenesis of transmembrane segment 7 of the *GLUT1* glucose transporter. *J. Biol. Chem.* 274:36176–36180.

Hruz, P.W., Mueckler, M.M. 2001. Structural analysis of the *GLUT1* facilitative glucose transporter. Mol. Membr. Biol. 18:183-193.

Hubbard, N.L., Pharr, D.M., Huber, S.C. 1991. Sucrose phosphate synthase and other sucrose metabolizing enzymes in fruit of various species. Physiol. Plant 82:191-196.

Imlau, A., Truernit, E., Sauer, N. 1999. Cell-to-cell and long-distance trafficking of the green fluorescent protein in the phloem and symplastic unloading of the protein into sink tissues. Plant Cell 11:309-322.

Ishiwatari, Y., Honda, C., Kawashima, I., Nakamura, S., Hirano, H., Mori, S., Fujiwara, T., Hayashi, H., Chino, M. 1995. Thioredoxin h is one of the major proteins in rice phloem sap. Planta 195:456-463.

Jeuffroy, M., Warembourg, F. R. 1991. Carbon transfer and partitioning between vegetative and reproductive organs in *Pisum sativum* L. Plant Physiol. 97,440-448.

Joshi, C.P. 1987. An inspection of the domain between putative TATA box and translation start site in 79 plant genes. Nucleic Acids Res 15: 6643–6653

Kahsay, R., Liao, L., Gao, G. 2005. An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. Bioinformatics. 21:1853-1858.

King, S.P., Lunn, J.E., Furbank, R.T. 1997. Carbohydrate content and enzyme metabolism in developing canola siliques. Plant Physiol. 114: 153-160

Kneller, E.L.P., Rakotondrafara, A.M., Miller, W.A. 2006. Cap-independent translation of plant viral RNAs. VirusRes 119:63-75.

Komor, E. Tanner, W. 1974. The Hexose-Proton Cotransport System of *Chlorella*: pH-Dependent Change in K_m Values and Translocation Constants of the Uptake System. J.Gen.Physiol. 64:568-581.

Kooiker, M., Airoldi, C.A., Losa, A., Manzotti, P.S., Finzi, L., Kater, M.M., Colombo, L. 2005. BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic *Arabidopsis* gene *SEEDSTICK*. The Plant Cell 17:722-729.

Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J. Mol. Biol. 305: 567-80.

Kulwithit, K. 1993. Fruit development and postharvest physiology of 'Pakchong 1' papaya. MS Thesis, Kasetsart University, Bangkok. (In Thai with English abstract).

Lazan, H., Ali, Z.M., Liang, K.S., Yee, K.L. 1989. Polygalacturonase activity and variation in ripening of papaya fruit with tissue depth and heat treatment. Physiol. Plantarum. 77: 93-97.

Lazan, H., Kasim, M., Ali, Z.M. 1995. β-Galactosidase, polygalacturonase and pectine esterase in differential softening and cell wall modification during papaya fruit ripening. Physiol. Plantarum. 95:106-112.

Lu¨tcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., Scheele, G.A. 1987. Selection of AUG initiation codons differs in plants and animals. EMBO J. 6:43–48.

Ma, T., Liu, Q., Li, Z., Zhang, X. 2002. Tonoplast H^+-ATPase in response to salt stress in *Populus euphratica* cell suspensions. *Plant Science* 163:499-505.

MacLeod, A.J., Pieris, N.M. 1983 Volatile components of papaya (*Carica papaya* L.) with particular reference to glucosinolate products. Journal of Agricultural and Food Chemistry. 31:1005-1008.

Maeshima, M., Yoshida, S. 1989. Purification and properties of vacuolar membrane proton-translocating inorganic pyrophosphatase from mung bean. J. Biol. Chem. 264:20068-20073

Marger, M.D., Saier, M.H. 1993. A major superfamily of trans-membrane facilitators that catalyze uniport, symport and antiport. Trends Biochem. Sci. 18:13-20.

Mason, M.G., Botella, J.R. 1997. Identification and characterisation of two 1-aminocyclopropane-1-carboxylate (ACC) synthase cDNAs expressed during papaya (*Carica papaya*) fruit ripening. *Aust. J. Plant Physiol.* 24:239-244

Mattheis,J.P., Fellman, J.K. 1999. Preharvest factors influencing flavor of fresh fruit and

vegetables. Postharvest Biol. Technology. 15:227-232.

M'Batchi, B., Delrot, S. 1984. Parachloromercuribenzenesulfonic acid - A potential tool for differential labeling of the sucrose transporter. Plant Physiol. 75:154-160.

McCollum, T.G., Huber, D.J., Cantliffe, D.J. 1988. Soluble sugar accumulation and activity of related enzymes during muskmelon fruit development. J. Amer. Soc. Hort. Sci. 113:399-403.

Miller, W.A., Waterhouse, P.M., Brown, J.W.S., Browning, K.S. 2001. The RNA World in Plants: Post-Transcriptional Control III. Plant Cell. 13:1710-1717.

Miron, D., Petreikov, M., Carmi, N., Shen, S., Levin, I., Granot, D., Zamski, E., Schaffer, A.A. 2002. Sucrose uptake, invertase localization and gene expression in developing fruit of *Lycopersicon esculentum* and the sucrose-accumulating *Lycopersicon hirsutum.* Physiol. Plant. 115: 35–47.

Monden. I., Olsowski, A., Krause, G., Keller, K. 2001. The Large Cytoplasmic Loop of the Glucose Transporter *GLUT1* Is an Essential Structural Element for Function. Biol. Chem. 382:1431-6730.

Mount,S.M. 1982. A catalogue of splice junction sequences. Nucl. Acids Res. 10:459-472

Muda, P., Angeles, D.E., Raveendranathan, P., Kosittrakun, M. 1994. Fruit growth and development. P. 35-47. *In* Rohani, M.Y. (ed.). Papaya-Fruit Development, Postharvest Physiology, Handling and Marketing in ASEAN. ASEAN Food Handling Bureau. Malaysia.

Nakasone, H.Y. 1986. Papaya. P. 277-301. In Monoelise, S.P. (ed.) CRC Handbook of Fruit Set and Development, CRC press, Boca Raton, Florida.

Nakasone, H.Y., Paull, R.E. 1998. Tropical Fruits, Chapter 10 Papaya. P239-269. CAB international.

Ong, H.T. 1983. Abortion during the floral-fruit development In *Carica papaya* in Serdang, Malasia. Pertanika. 6:105-107.

Oparka, K. J. 1990. What Is Phloem Unloading? Plant Physiol. 94:393-396.

Oparka, K.J., van Bel, A.J.E. 1992. Pathways of phloem loading and unloading: a plea for a uniform terminology. P249-254. In: Pollaock, C.J., Farrar, J.F., Gordon, A.J. (eds.). Carbon partitioning within and between organisms. BIOS Scientific, Oxford.

Oparka, K. J., Roberts, A. G., Boevink, P., Santa Cruz, S., Roberts, I., Pradel, K. S., Imlau, A., Kotlizky, G., Sauer, N., Epel, B. 1999. Simple, but not branched, plasmodesmata allow the nonspecific trafficking of proteins in developing tobacco leaves. Cell. 97:743-754.

Palmer, J.W., Cai, Y.L., Edjamo, Y. 1991. Effect of part-tree flower thinning on fruiting, vegetative growth and leaf photosynthesis in 'Cox's Orange Pippin' apple. J. Hort. Sci. 66:319-325.

Pantastico, Er.B. 1975. Preharvest factors affecting quality and physiology after harvest. P25-40. In: Pantastico, Er.B. (ed.). Postharvest physiology, handling and utilization of tropical and subtropical fruits and vegetables. AVI Pub., Westport, CT.

Parent, L. Supplisson, S., Loo, D.D.F., Wright, E.M. 1992a. Electrogenic properties of the cloned Na^+/glucose cotransporter: I. Voltage-clamp studies. J. Memb. Biol. 125:49-62.

Parent, L. Supplisson, S., Loo, D.D.F., Wright, E.M. 1992b. Electrogenic properties of the cloned Na^+/glucose cotransporter: II. A transport model under nonrapid equilibrium conditions. J. Memb. Biol. 125:63-79.

Patrick, J. W. 1990. Sieve element unloading: cellular pathway, mechanism and control. Physiol. Plant. 78:298-308.

Patrick, J. W. 1997. Phloem unloading: sieve element unloading and post-sieve element transport. Annu. Rev. Plant Physiol. Plant Mol. Biol. 48:191-222.

Patrick, J. W., Offler, C. E., Wang, X. D. 1995. Cellular pathway of photosynthate transport in coats of developing seed of *Vicia faba* L. and *Phaseolus vulgaris* L. 1. Extent of transport through the coat symplast. J. Exp. Bot. 46:35-47.

Paull, R.E. 1993. Pineapple and papaya. *In* Seymour, G., Taylor, J. and Tucker, G. (eds) Biochemistry of Fruit Ripening. Chapman & Hall, London, pp. 291-323.

Paull, R.E., Chen, N.J. 1983. Postharvest variation in cell wall-degrading enzymes of papaya (*Carica papaya* L.) during fruit ripening. Plant Physiol. 72: 382-385.

Pauli, R.E., Nishijima, W., Reyes, M., Cavaletto, C. 1997. A review of postharvest handling and losses during marketing of papaya (*Carica papaya* L.). Postharvest. Biol. Technol. 11:165-179.

Pavel, E.W., DeJong, T.M. 1993. Source- and sink-limited growth periods of developing peach

fruits indicated by relative growth analysis. J. Amer. Soc. Hort. Sci. 118:820-824.

Poll, L., Rindom, A., Toldam-Anderson, T.B., Hansen, P. 1996. Availability of assimilates and formation of aroma compounds in apples as affected by the fruit:leaf ratio. Physiol. Plant. 97:223–227.

Qui, Y.X., Nishima, M.S., Paull, R.E. 1995. Papaya fruit growth, calcium uptake, and fruit ripening. J. Amer. Soc. Hort. Sci. 120:246-253

Rausch, T. 1991. The hexose transporters at the plasma membrane and tonoplast of higher plants. Plant Physiol. 82:134-142.

Rohani, M.Y. 1994. Papaya-Fruit Development, Postharvest Physiology, Handling and Marketing in ASEAN. ASEAN Food Handling Bureau. Malaysia.

Roitsch, T., Tanner, W. 1994. Expression of a sugar-transporter gene family in a photoautotrophic suspension culture of *Chenopodium rubrum* L. Planta. 193:365-371.

Ruan, Y.L., Patrick J.W. 1995. The cellular pathway of postphloem sugar transport in developing tomato fruit. Planta 196:434-444

Ruan, Y.L., Patrick, J.W., Brady, C. 1997. Protoplast hexose carrier activity is a determinate of genotypic difference in hexose storage in tomato fruit. Plant, Cell and Environment. 20:341-349.

Sambrook, J., Fritsch, E.F., Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sankat, C.K., Maharaj, R. 1997. Papaya. P.167-208 In Mitra, S. (ed.) Postharvest Physiology and Storage of Tropical and Subtropical Fruits. CAB International. Oxon, UK.

Sauer, N., Baier, K., Gahrtz, M., Stadler, R., Stolz, J., Truernit, E. 1994. Sugar transport across the plasma membrane of higher plants. Plant Mol. Biol. 26:1671-1679.

Sauer, N., Caspari, T., Klebl, F., Tanner, W. 1990a. Functional expression of the *Chlorella* hexose transporter in *Schizosaccharomyces pombe*. Proc.Natl.Acad.Sci. USA 87:7949-7952.

Sauer, N., Friedländer, K., Gräml-Wicke, U. 1990b. Primary structure, genetic organization and heterogeneous expression of a glucose transporter from *Arabidopsis thaliana*. EMBO J.

9:3045-3050.

Sauer, N., Stadler, R. 1993. A sink specific H^+/monosaccharide co-transporter from *Nicotiana tabacum*: Cloning and heterologous expression in baker's yeast. Plant J. 4:601-610.

Sauer,N., Tanner, W. 1989. The hexose carrier from Chlorella cDNA cloning of a eucaryotic H^+-cotransporter. FEBS Letters. 259:43-46.

Sauer, N., Tanner, W. 1993. Molecular biology of sugar transporters in plants. Bot. Acta. 106:277-286.

Schulz, A. 1995. Plasmodesmal widening accompanies the short-term increase in symplasmic phloem unloading in pea root tips under osmotic stress. Protoplasma. 188:22-37.

Schuster, G., Lisitsky, I., Klaff, P. 1999. Polyadenylation and degradation of mRNA in the chloroplast. Plant Physiol. 120:937-944.

Seatter, M.J., De la Rue, S.A., Porter, L.M., Gould, G.W. 1998. QLS motif in transmembrane helix VII of the glucose transporter family interacts with the C-1 position of D-glucose and is involved in substrate selection at the exofacial binding site. Biochemistry. 37:1322-1326.

Selvaraj, Y., Pal, D.K., Subramanyam, M.D., Iyer, C.P.A. 1982. Changes in the chemical composition of four cultivars of papaya (*Carica papaya* L.) during growth and development. J. Hort. Sci. 57:135-143.

Shanmugavelu, K.G., Madhava Rao, V.N., Srinivasan, C. 1973. Studies on the effects of certain plant regulators and boron on papaya (*Carica papaya* L.). South Indian Hort. 21:19-26.

Sherson, S.M., Hemmann, G., Wallace, G., Forbes, S., Germain, V., Stadler, R., Bechtold, N., Sauer, N., Smith, S.M. 2000. Monosaccharide/proton symporter *AtSTP1* plays a major role in uptake and response of *Arabidopsis* seeds and seedlings to sugars. Plant J. 24:849-857.

Shiratake, K., Kanayama, Y., Yamaki, S. 1997. Characterization of hexose transporter for facilitated diffusion of the tonoplast vesicles from pear fruit. Plant Cell Physiol. 38(8):910-916.

Singh, A. 1980. Fruit Physiology and Production. Kalyani Publ., New Delhi.

Sjölund, R.D.1997. The phloem sieve element- A river runs through it. Plant Cell. 9:1137-1146.

Stadler, R., Lauterbach, C., Sauer, N. 2005a. Cell-to-cell movement of green fluorescent protein reveals post-phloem transport in the outer integument and identifies symplastic domain in

Arabidopsis seeds and embryos. Plant Physiol.

Stadler, R., Wolf, K., Hilgarth, C., Tanner, W., Sauer, N. 1995. Subcellular Localization of the Inducible Chlorella *HUP1* Monosaccharide-H+ Symporter and Cloning of a Co-Induced Galactose-H+ Symporter. Plant Physiol. 107:33-41.

Stadler, R., Wright, K. M., Lauterbach, C., Amon, G., Gahrtz, M., Feuerstein, A., Oparka, K.J., Sauer, N. 2005b. Expression of GFP-fusions in Arabidopsis companion cells reveals non-specific protein trafficking into sieve elements and identifies a novel post-phloem domain in roots. Plant J. 41:319-331.

Stolz, J., Stadler, R., Opekarova, M., Sauer, N. 1994. Functional reconstitution of the solubilied *Arabidopsis thaliana STP1* monosaccharide-H^+symporter in lipid vesicles and purification of the histidine tagged protein from transgenic *Saccharomyces cerevisiae*. Plant J. 6:225-233.

Sung, S.J., Xu, D.P., Black, C.C. 1988. A reassessment of glycolysis and gluconeogenesis in higher plants. Physiol. Plant. 72:650-654.

Taiz, L, Zeiger, E. 2002. *Plant Physiology, Third edition*. 690p. Sinauer Associates, Inc., MA.

Tee, E.S., Mohd. Ismail, N., Mohd. Nasir, A., Khatijah, I. 1988. Nutrient composition of Malaysian foods. ASEAN Food Habits Project, *Cited by* Rohani, M.Y. 1994. Papaya-Fruit Development, Postharvest Physiology, Handling and Marketing in ASEAN. ASEAN Food Handling Bureau. Malaysia.

Terrier N, Glissant D, Grimplet, J. Barrieu, F., Abbal, P., Couture, C., Ageorges, A., Atanassova, A., Le´on, C., Renaudin, J., De´dalde´champ, F., Romieu, C., Delrot, S., Hamdi, S. 2005. Isogene specific oligo arrays reveal multifaceted changes in gene expression during grape berry (*Vitis vinifera* L.) development. Planta 222: 832–847.

Thorne, J. H. 1985. Phloem unloading of C and N assimilates in developing seeds. Annu. Rev. Plant Physiol. 36:317-343.

Thorsch, J., Esau, K. 1981. Nuclear degeneration and the association of endoplasmic reticulum with the nuclear envelope and microtubules in maturing sieve elements of *Gossypium hirsutum*. J. Ultrastruct. Res. 74, 195-204.

Truernit, E., Schmid, J., Epple, P., Illig, J., Sauer, N. 1996. The sink-specific and stress regulated

Arabidopsis STP4 gene: enhanced expression of a gene encoding a monosaccharide transporter by wounding, elicitors and pathogen challenge. Plant Cell. 8:2169-2182.

Truernit, E., Stadler, R., Baier, K., Sauer, N. 1999. A male gametophyte-specific monosaccharide transporter in *Arabidopsis*. Plant J. 17:191-201.

Wang, F., Smith, A.G., Brenner, M.L. 1994. Temporal and spatial expression pattern of sucrose synthase during tomato fruit development. Plant Physiol. 104:535-540.

Weber, H., Borisjuk, L., Heim, U., Buchner, P., Wobus, U. 1995. Seed coat-associated invertases of faba bean control both unloading and storage functions: cloning of cDNAs and cell type specific expression. Plant Cell. 7:1835-1846.

Weber, H., Borisjuk, L., Heim, U., Sauer, N., Wobus, U. 1997. A role for sugar transporters during seed development: Molecular characterization of a hexose and a sucrose carrier in Fava bean seeds. Plant Cell. 9:895-908.

Weig, A., Franz, J., Sauer, N., Komor, E. 1994. Isolation of a family of cDNA clones from *Ricinus communis* L. with close homologyto the hexose carriers. J. Plant Physiol. 143:178-183.

Weschke, W., Panitz, R., Gubatz, S., Wang, Q., Radchuk, R., Weber, H., Wobus, U. 2003. The role of invertases and hexose transporters in controlling sugar ratios in maternal and filiar tissues of barley caryopses during early development. Plant J. 33:395-411

Williams, L.E., Lemoine, R., Sauer, N. 2000. Sugar transporters in higher plants - a diversity of roles and complex regulation. Trends in Plant Science. 5:283-290

Yamaki, S., Asakura, T. 1988. Energy coupled transport of sorbitol and other sugars into the protoplast isolated from apple fruit flesh. Plant Cell Physiol. 29:961-967.

Yee, W., Akamine, E.K. and Aoki, G. M. 1970. Papayas in Hawaii. University of Hawaii Cooperative Extension service. Circular 436.

Zhang, D.P., Lu, Y.M., Wang, Y.Z., Duan, C.Q., Yan, H.Y. 2001. Acid invertase is predominantly localized to cell walls of both the practically symplastically isolated sieve element/companion cell complex and parenchyma cells in developing apple fruits. Plant, Cell and Environment 24:691–702.

Zhang, X. Y., Wang, X. L., Wang, X. F., Xia, G. H., Pan, Q. H., Fan, R. C., Wu, F. Q., Yu, X. C.,

Zhang, D. P. 2006. A Shift of phloem unloading from symplasmic to apoplastic pathway is involved in developmental onset of ripening in grape berry. Plant Physiol. 142:220-232.

Zhou, L. 1999. Source-sink relationship during papaya fruit development and ripening. Ph.D. dissertation, University of Hawaii at Manoa, 146p.

Zhou, L., Christopher, D.A., Paull, R.E. 2000. Defoliation and fruit removal effects on papaya fruit production, sugar accumulation, and sucrose metabolism. J. Amer. Soc. Hort. Sci. 125:644-652.

Zhou, L., Pauli, R.E. 2001. Sucrose metabolism during papaya (*Carica papaya*) fruit growth and ripening. J. Amer. Soc. Hort. Sci.126:351-357

Zhou, L., Chen, C.C., Ming, R., Christopher, D.A., Paull, R.E.. 2003. Apoplastic invertase and its enhanced expression and post translation control during fruit maturation and ripening. J. Amer. Soc. Hort. Sci.128:628-635.

Ziegler, H. 1975. Nature of transported substances. *In* JA Milburn, MH Zimmermann, eds, Encyclopedia of plant physiology, Vol 1. Springer Verlag, Berlin, pp 59-100.